National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

Molecular Analysis of Human tRNA Genes

by

Ronald W. Henry

A Thesis

Submitted to the Faculty of Graduate Studies and Research

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Department of Microbiology

Edmonton, Alberta

Spring, 1992

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Service | Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-73222-9

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

Name of Author: Ronald W. Henry

Title of Thesis: Molecular Analysis of Human tRNA Genes

Degree: Doctor of Philosophy

Year this Degree Granted: 1992

Permission is hereby granted to the University of Alberta to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed without the author's permission.

Bill Henry

(Ronald William Henry)

4313 - 47 Street
Red Deer, Alberta
Canada. T4N 1P6

Date: April 10/92

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled, Molecular Analysis of Human tRNA Genes, submitted by Ronald W. Henry, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Kenneth Roy, supervisor

Dr. Susan Jensen

Dr. Heather McDermid

Dr. John Bell

Dr. Michael Pickard

for Dr. Gordon Tener, External

Date: April 1/92

ABSTRACT

The chromosomal locations of four human tRNA gene-containing DNA fragments have been determined. Two fragments, containing at least four $tRNA^{Tyr}$ genes, have been mapped to chromosome 14. Another fragment, containing a single $tRNA^{Gly}$ gene, has been mapped to chromosome 1, while a fourth fragment, containing two $tRNA^{Phe}$ genes and two $tRNA^{Lys}$ genes, has been located on chromosome 11. This is the first example of human tRNA genes having been located on chromosome 11. Additionally, no other chromosomal locations for human $tRNA^{Phe}$ genes had been previously known.

In order to accelerate the recovery of human tRNA genes for further studies, four different PCR strategies were used to amplify tRNA genes directly from the human genome. Five DNA fragments containing ten tRNA genes or pseudo-genes for the phenylalanine and tyrosine tRNA gene families were isolated by using "tRNA PCR", where PCR is performed using tRNA gene-specific primers which are directed at the 5'- and/or 3'-conserved gene sequence(s), and are oriented through the gene. Flanking sequence comparisons of three $tRNA^{Phe}$ gene-containing fragments revealed no similarities to other previously characterized $tRNA^{Phe}$ gene flanking sequences. Two fragments, each containing two tandemly repeated human $tRNA^{Tyr}$ genes, show greater than 95% homology to a DNA fragment isolated from a human-λ recombinant library. At least two fragments containing $tRNA^{Phe}$ genes and flanking sequences were amplified from the human genome using "Inverted PCR". In this technique genomic DNA was first cleaved with a restriction endonuclease and the fragments ligated under conditions that favor circularization. PCR was performed using primers that are arranged in an opposite orientation to that used for tRNA PCR. Amplification of fragments was also observed from inverted PCR for the $tRNA^{Trp}$ and $tRNA^{His}$ gene families. Specific amplification of tRNA gene-containing fragments was observed for "Alu-tRNA PCR" and "Linker PCR", when performed as a two step PCR using

biotinylated tRNA gene-specific primers in conjunction with streptavidin-agarose selection of tRNA gene sequences. The second primer used in Alu-tRNA PCR was complementary to a region of the human Alu repeat sequence, whereas in Linker PCR, the second primer binding site was provided by the ligation of linkers to the termini of DNA fragments generated by restriction endonuclease digestion of human genomic DNA. By utilizing these two methods, not only is it possible to isolate tRNA gene sequences, but it is also possible to "walk" from the tRNA gene to the flanking regions by using Linker PCR. These results suggest that tRNA PCR will be very useful for isolating a moderate proportion of human genes for some tRNA gene families where the tRNA genes are clustered, as well as revealing the presence of intron sequences in some tRNA genes. Inverted PCR is likely to be the most useful for the isolation of tRNA genes and their flanking sequences. This method should allow the isolation of most, if not all, members of a tRNA gene family, regardless of genomic organization.

Acknowledgement

I would like to thank Drs. Kenneth Roy, Jamie Doran, and Xiaoning Wu for their patience and understanding during my transition to the world of science. Their help and insight was invaluable. My lab mates and friends, Domenic Spadafora and Donald Netolitzky, deserve special mention for putting up with me, day in and day out, and also for doing their share of the work around here. Those people involved in the "best summer(s) ever" also deserve something. To Greg, Dylan, Rundle, Rich, Steve, Jeff, Ange and Tom: may you bat at least .500 in your future endeavors. My deepest appreciation and love goes to my parents, my sister Heather, Lane McNaughton, and Nicole Trepanier for their support and encouragement during this period. Without them I may not have taken this road less travelled.

Table of Contents

	page
Library Release Form	i
Title Page	ii
Committee Approval Page	iii
Abstract	iv
Acknowledgements	vi
Table of Contents	vii
List of Tables	x
List of Figures	xi
List of Abbreviations	xiv
1 Introduction	1
1.1 The function and structure of tRNA	1
1.2 The structure and transcription of tRNA genes	8
1.3 The organization of the human genome	12
1.4 The genomic organization of human tRNA genes	15
1.5 The Polymerase Chain Reaction	19
1.5.1 tRNA PCR	22
1.5.2 Alu-tRNA PCR	25
1.5.3 Inverted PCR	26
1.5.4 Linker PCR	32
1.6 Objectives of this study	35
2. Methods and Materials	38
2.1 Laboratory materials and enzymes	38
2.2 Isolation of human and mouse genomic DNA	38
2.3 Preparation of DNA probes	39

2.4 Human tRNA gene localization studies..... 40

2.4.1 Standardization of hybridization conditions..... 40

A. Hybridization of oligonucleotides to human genomic DNA..... 40

B. Effect of temperature on hybridization of an M13 primer extension probe to human genomic DNA..... 40

C. Effect of temperature on hybridization of random primer probes to human genomic DNA..... 41

D. Effect of probe length..... 41

E. Hybridization of tRNA gene-containing random primer probes to mouse and human genomic DNA: probe descriptions and hybridization conditions..... 42

2.4.2 Hybridization of tRNA gene containing probes to mouse/human somatic cell hybrid chromosomal mapping panels..... 43

2.5 Streptavidin-agarose selection of PCR amplification products..... 44

2.6 Isolation of human tRNA gene sequences..... 44

2.6.1 tRNA PCR..... 44

2.6.2 Inverted PCR..... 50

2.6.3 Alu-tRNA PCR..... 54

2.6.4 Linker PCR..... 55

2.7 Cloning and screening of PCR amplification products..... 57

2.8 DNA sequencing..... 58

3. Results..... 59

3.1 Standardization of hybridization conditions..... 59

3.1.1 Oligonucleotide hybridization..... 59

3.1.2 M13 primer extension..... 59

3.1.3 Random primer extension..... 60

3.1.4 Hybridization to mouse and human genomic DNA..... 68

3.2 Human tRNA gene chromosomal localization studies..... 83

3.3 Isolation of human tRNA genes..... 98
3.3.1 tRNA PCR..... 98
A. $tRNA^{Tyr}$..... 98
B. $tRNA^{Phe}$..... 109
C. $tRNA^{Lys}$..... 123
D. $tRNA^{Arg}$..... 123
E. $tRNA^{Trp}$..... 128
F. $tRNA^{Met}$..... 128
3.3.2 Inverted PCR..... 133
A. $tRNA^{Phe}$..... 133
B. $tRNA^{His}$..... 147
C. $tRNA^{Trp}$..... 152
3.3.3 Alu-tRNA PCR..... 152
A. $tRNA^{Lys}$..... 152
B. $tRNA^{Phe}$..... 157
3.3.4 Linker PCR - $tRNA^{Phe}$..... 160
4. Discussion..... 166
4.1 Standardization of hybridization conditions..... 166
4.2 Human tRNA gene chromosomal mapping studies..... 172
4.3 Isolation of human tRNA genes..... 175
4.3.1 tRNA PCR..... 176
4.3.2 Alu-tRNA PCR..... 185
4.3.3 Inverted PCR..... 187
4.3.4 Linker PCR..... 194
4.4 Conclusions..... 197
5. References..... 200

List of Tables

		page
Table 1	Summary for hybridization of tRNA gene-containing probes to a human chromosomal mapping panel	94
Table 2	Percentage discordancy table for hybridization of tRNA gene-containing probes to a human chromosomal mapping panel	96
Table 3	Summary of known chromosomal locations for human tRNA gene loci	97

List of Figures

		page
Figure 1	The secondary "cloverleaf" stucture of tRNA indicating the standard tRNA numbering notation	3
Figure 2	The folding pattern involved in tertiary interactions for tRNAs	7
Figure 3	Schematic representation of the Polymerase Chain Reaction	21
Figure 4	Schematic representation of "tRNA PCR"	24
Figure 5	Schematic representation of "Alu-tRNA PCR"	28
Figure 6	Schematic representation of "Inverted PCR"	30
Figure 7	Schematic representation of "Linker PCR"	34
Figure 8	tRNA PCR primer diagram	47
Figure 9	Inverted PCR primer diagram	52
Figure 10	Hybridization of an M13 primer extension probe (specific for a $tRNA^{Lys}$ gene loci) to human genomic DNA	62
Figure 11	Hybridization of a human $tRNA^{Gln}$ gene probe to human genomic DNA	64
Figure 12	Effect of flanking sequence on hybridization of $tRNA^{Gln}$ specific probes to human genomic DNA	67
Figure 13	Hybridization of a λHtM2 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA	70
Figure 14	Hybridization of a λHtM4 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA	72
Figure 15	Hybridization of a λHtM6 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA	74
Figure 16	Hybridization of a λHt4 ($tRNA^{Gln}$, $tRNA^{Leu}$, $tRNA^{Lys}$) specific probe to human and mouse genomic DNA	76
Figure 17	Hybridization of a λHt7 ($tRNA^{Lys}$, $tRNA^{Phe}$) specific probe to human and mouse genomic DNA	78
Figure 18	Hybridization of a λHt8 ($tRNA^{Gly}$) specific probe to human and mouse genomic DNA	80
Figure 19	Hybridization of a λHt9 ($tRNA^{Gly}$) specific probe to human and mouse genomic DNA	82

Figure 20	Hybridization of a λHtM6 ($tRNA^{Tyr}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel	85
Figure 21	Hybridization of a λHtM2 ($tRNA^{Tyr}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel	87
Figure 22	Hybridization of a λHt7 ($tRNA^{Phe}$, $tRNA^{Lys}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel	89
Figure 23	Hybridization of a λHt9 ($tRNA^{Gly}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel	91
Figure 24	tRNA PCR for the human $tRNA^{Tyr}$ gene family	100
Figure 25	Schematic representation of amplification products from tRNA PCR using primers specific for the human $tRNA^{Tyr}$ gene family	102
Figure 26	Nucleotide sequence comparisons of $tRNA^{Tyr}$ gene-containing recombinants pHy6, pHy9, pHy10, pHy15 and pHy20	104
Figure 27	Amplification products from PCR using various $tRNA^{Tyr}$ intron-specific primers and a primer common to the intergenic flanking sequence found within pHy6, pHy10, pHy15 and pHy20	108
Figure 28	Nucleotide sequence comparisons of $tRNA^{Tyr}$ gene intron sequences	111
Figure 29	tRNA PCR using primers specific for the human $tRNA^{Phe}$ gene family	113
Figure 30	Schematic representation of amplification products from tRNA PCR for the $tRNA^{Phe}$ gene family	116
Figure 31	Nucleotide sequence of a human $tRNA^{Phe}$ gene and flanking sequence contained within pHf1-39	118
Figure 32	Nucleotide sequence of the $tRNA^{Phe}$ gene and pseudogenes within pHf7-30 and pHf7-31	121
Figure 33	tRNA PCR using primers specific for the human $tRNA^{Lys}$ gene family	125
Figure 34	tRNA PCR using primers specific for the human $tRNA^{Arg}$ gene family	127
Figure 35	tRNA PCR using primers specific for the human $tRNA^{Trp}$ gene family	130
Figure 36	tRNA PCR using primers specific for the human $tRNA^{Met}$ gene family	132

Figure 37	Inverted PCR for human genomic DNA digested with *Cfo*I, *Hin*PI, or *Hae*II using primers specific for the human $tRNA^{Phe}$ gene family	135
Figure 38	Nucleotide sequence of the $tRNA^{Phe}$ genes within pCfo1, pCfo2, and pCfo35	138
Figure 39	Nucleotide sequence comparisons of the $tRNA^{Phe}$ gene-containing recombinants pHae2, pHae19, pHae20 and pHae23	140
Figure 40	Nucleotide sequence of the $tRNA^{Phe}$ genes or pseudogenes within pHnp2, pHnp4, pHnp11, and pHnp12	142
Figure 41	Amplification products from PCR from human genomic DNA using primers that are complementary to DNA regions flanking the $tRNA^{Phe}$ genes found in pHnp4 and pHnp12	146
Figure 42	Inverted PCR for human genomic DNA digested with *Hin*dIII using primers specific for the human $tRNA^{Phe}$ gene family	149
Figure 43	Inverted PCR for human genomic DNA digested with *Hpa*II or *Hin*dIII using primers specific for the human $tRNA^{His}$ gene family	151
Figure 44	Inverted PCR for human genomic DNA digested with *Hpa*II or *Hin*dIII using primers specific for the human $tRNA^{Trp}$ gene family	154
Figure 45	Alu-tRNA PCR using primers specific for the human $tRNA^{Lys}$ gene family	156
Figure 46	Alu-tRNA PCR using primers specific for the human $tRNA^{Phe}$ gene family	159
Figure 47	Linker PCR for the human $tRNA^{Phe}$ gene family	162
Figure 48	Linker PCR for the human $tRNA^{Phe}$ gene family using nested primers	165
Figure 49	Schematic representation of the generation of artifact sequences by tRNA PCR	178
Figure 50	Schematic representation of the generation of artifacts by inverted PCR	191
Figure 51	Schematic representation of the generation of artifacts by linker PCR	195

List of Abbreviations

Amp	ampicillin
bp	base pair(s)
IPTG	isopropyl-β-thiogalactoside
kb	kilobase(s)
NaOAc	sodium acetate
nt	nucleotide(s)
Ψ	pseuoduridine
snRNA	small nuclear RNA
TFIID	transcription factor IID
TFIIIA	transcription factor IIIA
TFIIIB	transcription factor IIIB
TFIIIC	transcription factor IIIC
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1. INTRODUCTION

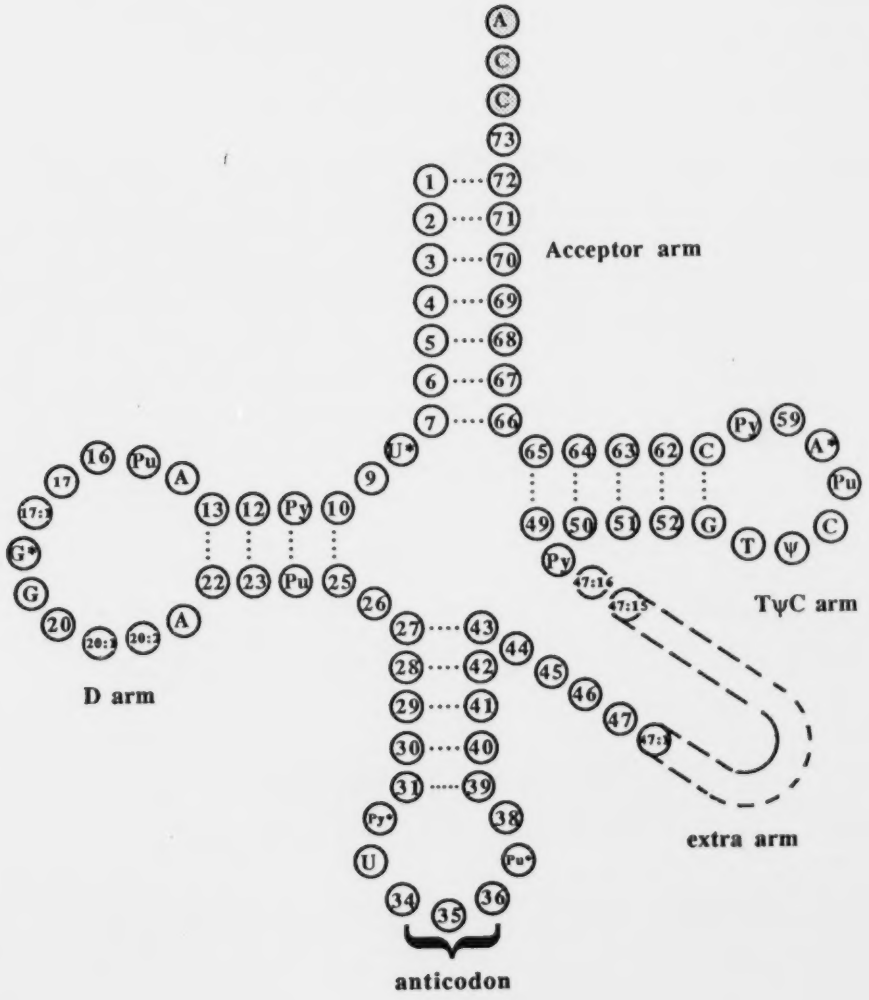
1.1 The function and structure of tRNA

In the translation of genetic information from nucleic acid sequence to protein sequence, transfer RNAs (tRNA) play a crucial intermediate role. Transfer RNAs are essential for the initiation of protein synthesis and for the elongation of the growing polypeptide chains. During protein synthesis, the amino acid sequence of the polypeptide is determined by the interaction of the tRNA and the mRNA on the ribosome. The tRNA component of this interaction has two crucial properties: (1) each tRNA can be aminoacylated with only one amino acid and, (2) each tRNA contains a trinucleotide sequence, the anticodon, which is complementary to the codon representing the esterified amino acid. The anticodon enables the tRNA to recognize the codon via complementary base pairing. All tRNAs capable of being aminoacylated with the same amino acid are referred to as isoaccepting tRNAs. Because they must function interchangeably on the ribosome during protein synthesis, all tRNAs have similar primary (Sprinzl et al., 1987), secondary (Holley et al., 1965) and tertiary (Kim et al., 1974, Klug et al., 1974) structures. In addition, the structure of tRNAs reflects the influence of interactions between the tRNA or the tRNA gene and various proteins including transcription factors, modifying enzymes, and the aminoacyl tRNA synthetases.

The familiar cloverleaf representation of the tRNA secondary structure (see Figure 1) highlights the five major arms of this molecule. The "D arm" is composed of a double stranded stem region (the D stem) composed of at least 3 (sometimes 4) bp and the D loop comprising an unpaired region containing at least 9 nucleotides. This arm is named for the presence of a highly conserved modified nucleotide, dihydrouridine, at position 18 and frequently other positions in this loop. The "anticodon arm" contains the anticodon stem and the anticodon loop. The stem is a double stranded helix containing 5 bp. The anticodon loop contains 7 unpaired

Figure 1 - The secondary "cloverleaf" structure of tRNA indicating the standard tRNA numbering notation.

For invariant bases, the actual base is indicated. Semi-invariant bases are shown as Py (pyrimidine) or Pu (purine). An * indicates those bases that are usually modified in the tRNA. Circles containing small numbers represent nucleotides which are not present in all tRNAs. The terminal CCA is usually added post-transcriptionally in eukaryotes but is gene encoded in prokaryotes. (Adapted from Kim et al., 1974)


Acceptor arm
D arm
TψC arm
extra arm
anticodon

nucleotides and is the region of the tRNA molecule that contains the anticodon. The "extra arm" or "variable arm" found between the anticodon stem and the TψC arm is composed of between 3 and 21 nucleotides. The number of bases present in this loop varies for different isoaccepting tRNAs. The "TψC arm" is named for the presence of this triplet sequence, containing a highly conserved nucleotide, pseudouridine, at position 55. The double stranded helical stem region contains 5 bp. The loop usually usually contains 7 bases. The "amino acid accepting stem" is a double stranded helix usually composed of at least seven base pairs plus an extra unpaired 3'-terminal nucleotide. One exception is the bovine histidine tRNA which has a 5' terminal G added post-transcriptionally (Sprinzl et al., 1987). Transfer RNAs typically have a CCA trinucleotide sequence attached to the 3'-terminus. In eukaryotes, this trinucleotide sequence is attached to the amino acid accepting arm post-transcriptionally, whereas in prokaryotes this sequence is gene encoded. The final functional molecule is achieved via the covalent attachment of the appropriate amino acid to the 3' end of the tRNA (reviewed in Sharp et al., 1985).

The process of charging the tRNA to form aminoacyl tRNA is catalyzed by a group of highly specific enzymes, the aminoacyl tRNA synthetases (AAS). There are at least 20 aminoacyl tRNA synthetases. Although there are some rare exceptions, there is normally a single AAS for each amino acid. An AAS must correctly recognize and activate an amino acid and then esterify it to each member of the cognate set of tRNAs. Thus, within this homogeneous tRNA framework variation must exist so that each tRNA is recognizable to its cognate AAS. The term "tRNA identity" has been coined to describe those features of a tRNA molecule which make that tRNA recognizable to one AAS and prevent its recognition by all other AASs. The approach that has proven the most useful in the establishment of tRNA identity is one in which an attempt is made to alter with the fewest possible changes, the specificity of aminoacylation of the tRNA. The accuracy of tRNA recognition by an AAS is a critical

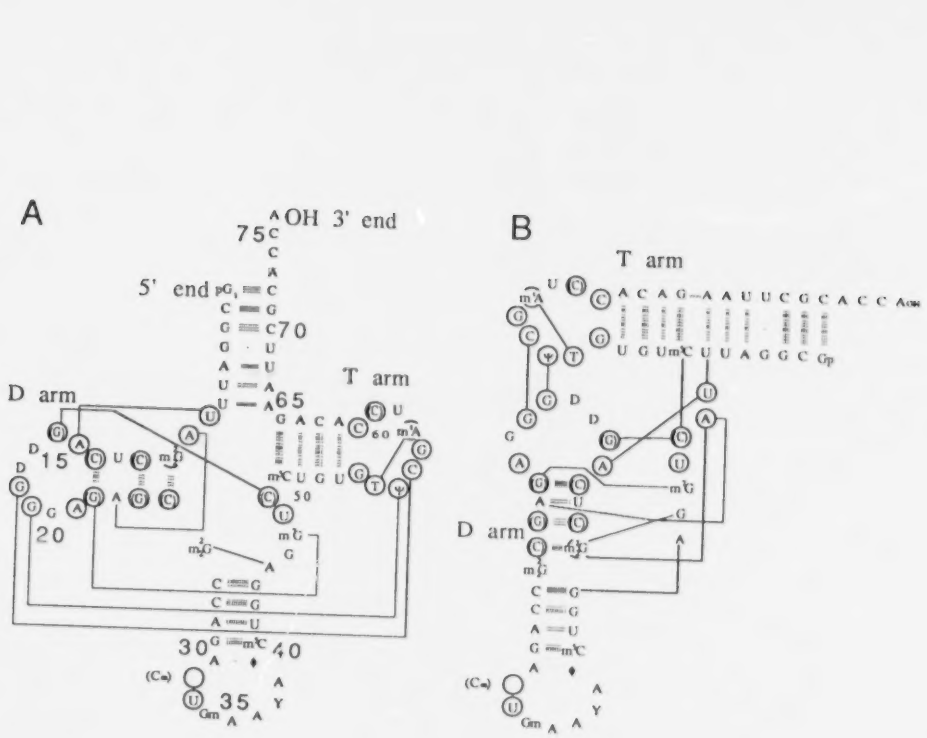
element in achieving the specificity of aminoacylation and ultimately this is crucial for fidelity in protein biosynthesis (reviewed in Normanly and Abelson 1989). The presence of an AAS and its cognate tRNA in proper balance is critical to maintaining this fidelity (Swanson et al., 1988).

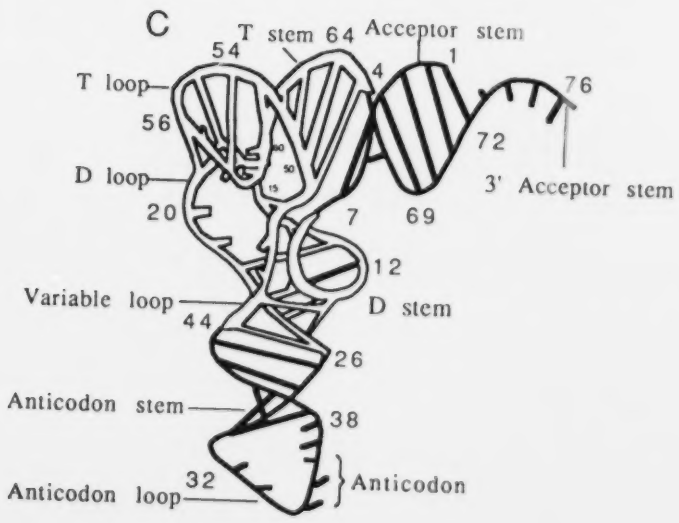
Intrinsic to the analysis of the tRNA-AAS interaction is a consideration of the tertiary structure of this tRNA-protein complex. After assuming the functional, complex, L-shaped tertiary structure (see Figure 2) the amino acid accepting arm of the tRNA is situated at one end of the molecule while the anticodon is effectively located at the other. The L structure is formed by folding the cloverleaf structure between the D and T stems. This tertiary structure is created and maintained largely by hydrogen bonds in the stem regions of the tRNA and an additional 9 tertiary bonds between bases which are unpaired in the secondary structure of the tRNA (Kim, 1978). Most of the invariant and semi-invariant bases of a tRNA are important in the formation of this tertiary structure (Clark, 1978). Three general areas of the tRNA molecule have been implicated in contacting the synthetase: (1) the acceptor stem is involved because it must be charged at the 3' terminus with the amino acid, (2) the D stem is implicated by reactions in which the AAS is cross-linked photochemically to the tRNAs and, (3) often the anticodon is also linked to the enzyme. A model has been proposed in which the tRNA synthetase binding occurs along the inside of the L-shaped tRNA molecule. This has been supported by X-ray studies of the cocrystals of *Escherichia coli* glutamine-tRNA synthetase with $tRNA^{Gln}$ (Rould et al., 1989; Rould et al., 1991) and also of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$ (Ruff et al., 1991). Once a tRNA has been charged, the amino acid plays no further role in its specificity, which is determined exclusively by the anticodon of the tRNA (Chapeville et al., 1962).

All AASs function by a two step mechanism. First, the amino acid reacts with ATP to form an aminoacyl adenylate, releasing pyrophosphate. The activated amino acid is then transferred to the tRNA forming an aminoacyl ester with the 2'- or 3'-OH

Figure 2 - The folding pattern involved in tertiary interactions for tRNAs.

Those bases that are unpaired when represented in the cloverleaf structure (A) are usually involved in base pairing with other regions of the tRNA (B) to aid in achieving the final L-shaped tertiary structure (C). The final tRNA tertiary structure, achieved by folding between the D arm and the T arm, is stabilized by hydrogen bonds and base stacking interactions. In this conformation, the anticodon is located at one end of the molecule while the amino acid acceptor stem is located at the other. (Adapted from Kim et al., 1974)


A
OH 3' end
5' end
D arm
T arm
B
T arm
D arm


C
T stem
Acceptor stem
T loop
D loop
3' Acceptor stem
Variable loop
D stem
Anticodon stem
Anticodon loop
Anticodon

of the tRNA, in the process releasing AMP. Despite the common mechanism of AAS action, the structures of the synthetases for different tRNA families vary widely (reviewed in Schimmel, 1987; Schimmel and Söll, 1979). Many early attempts to detect similarities in sequence between isoaccepting tRNAs or to induce chemical alterations that alter their charging had shown that the basis for tRNA recognition does not lie in some simple feature of the primary or secondary structure alone. The variability in behavior amongst tRNAs, however, suggests that the locations of the determinants for tRNA identity are idiosyncratic (Schulman and Abelson, 1988).

Additionally, tRNA molecules are among the most highly modified biologically active macromolecules known. Modification of these molecules may aid in achieving the correct configuration of the tRNA for recognition by its aminoacyl tRNA synthetase, or these modifications may be important for correct processing of the tRNA to assume its final functional structure. Deletion of the intron sequence from a human $tRNA^{Tyr}$ gene prevented the post-transcriptional modification of a uridine to pseudouridine found in all $tRNA^{Tyr}$ species at position 35 (van Tol and Beier, 1988). Some modifications are also important for specific binding to the codon of the mRNA (e.g. $tRNA^{Ile}_{CAT} \rightarrow ATA$). Sequences that are necessary for proper modification or function are likely to be highly conserved.

1.2 The structure and transcription of tRNA genes

The coding regions of tRNA genes vary in length from ~70 to 100 bp and contain certain common structural features. Each gene must code for the stem-loop elements and several other invariant and semi-invariant bases which contribute to the mature tRNA structure or function. Human tRNA genes usually do not contain intron sequences although some exceptions are known. The first example of intron sequences in human tRNA genes was reported by MacPherson and Roy (1986) when it was shown that $tRNA^{Tyr}$ genes encoded introns of either 20 or 21 bp in length. More

recently, Green et al. (1990) utilized the polymerase chain reaction to confirm the presence of introns in all genes coding for $tRNA^{Tyr}$. A novel 16 bp intron for $tRNA^{Tyr}$ and introns of 22-24 bp in length for some human $tRNA^{Leu}$ genes were also described.

In addition to the conservation of invariant bases for structural reasons, the coding regions of tRNA genes are also under constraints due to the presence of intragenic control regions (ICRs) which serve as promoters for transcription by RNA polymerase III (for reviews see Hall et al., 1982; Ciliberto et al., 1983; Sharp et al., 1985). Transcription in eukaryotic cells is performed by three different nuclear RNA polymerases; RNA polymerases I, II, and III. Each enzyme is responsible for transcribing a different class of genes. As well as transcribing tRNA genes, other genes transcribed by RNA polymerase III include 5S rRNA, U6 snRNA, 7SK, and H1 genes.

RNA polymerase III and its associated transcription factors may recognize very divergent promoter sequences. The ICR of 5S rRNA genes was shown to be located in a single region corresponding to the coordinates 55-80, designated the "box A" region, and in the contiguous sequence from coordinates 81-99, designated the "box C" region. Conversely, the dissection of tRNA gene ICRs by deletion mutation analysis demonstrated that they are discontinuous structures with essential "boxes" separated by non-essential nucleotides (DeFranco et al., 1980; Galli et al., 1981; Sharp et al., 1981; Ciliberto et al., 1982; Baker and Hall, 1984; Pieler et al., 1985). The "box A" sequence is found at coordinates 8-19 and is conserved among different RNA polymerase III transcribed genes. The box B region is the major quantitative determinant of promoter strength, however tRNA gene flanking sequences have also been shown to have a role in determining transcription rate (Arnold et al., 1986; Doran et al., 1987; reviewed in Sharp et al., 1985). Limited homology exists between conserved regions of the "box B" ICRs (coordinates 52-62) of tRNA genes and other

RNA polymerase III transcribed genes, including the Alu repeat sequence family in humans.

The promoter regions of the RNA polymerase III transcribed U6 snRNA genes demonstrate a radically different promoter arrangement when compared to either the 5S rRNA or tRNA genes. The promoter region for these genes is, in contrast, found as an extragenic control region consisting of a distal sequence element (DSE) at approximately position -220, a proximal sequence element (PSE) located around position -50, and a TATA box like region located at position approximately -25 relative to the start of the U6 snRNA gene. The human U2 snRNA genes, which are transcribed by RNA polymerase II, also exhibit the conserved DSE and PSE regions. Mutations in the TATA box of the U6 gene can convert transcription of this gene by RNA polymerase III to that of RNA polymerase II. Conversely, placing the U6 TATA box region upstream from the U2 genes converts these genes to RNA polymerase III transcribed units (Lobo et al., 1990). This diversity of promoter arrangements for RNA polymerase III transcribed genes may reflect the use of specialized transcription factors to control the transcription of these genes by the same RNA polymerase.

The RNA polymerase III transcription complexes of all eukaryotes share extensive similarities. These include the use of the ancillary transcription factors TFIIIA, TFIIIB and TFIIIC. The factors TFIIIA and TFIIIC from yeast and higher eukaryotes share similar general DNA binding properties and size (Pieler et al., 1987; Wang and Weil, 1989; Braun et al., 1989; Chalice and Segall, 1990). Similarly, the transcription factor TFIIIB preparations from humans and yeast contain a polypeptide of ~60,000 Da (Klekamp and Weil, 1986; Waldschmidt et al., 1988). In yeast, it has been shown that TFIIIB fails to bind DNA specifically in the absence of TFIIIC and these two factors are the last proteins to assemble on the DNA prior to the binding of RNA polymerase III (Lassar et al., 1983; Bieker et al., 1985; Segall, 1986). The binding of TFIIIB and TFIIIC is mediated by the binding of TFIIIA to the box A and

box C regions (e.g. 5S RNA genes). Conversely, the protein TFIIIC has been shown to bind specifically to the box A and box B regions (e.g. tRNA genes). Kassavetis et al. (1990) demonstrated that the yeast TFIIIA and TFIIIC function as assembly factors for TFIIIB and once they have properly positioned TFIIIB upstream of the transcription start sites of the 5S or tRNA genes they are dispensable for transcription by RNA polymerase III. The upstream DNA binding ability of TFIIIB is also mimicked by factors involved in the transcription of genes by RNA polymerases I and II. For RNA polymerase I, this factor may be pol I factor D (Sollner-Webb and Tower, 1986), whereas the analogous function may be served by TFIID for RNA polymerase II transcribed genes (Reinberg et al., 1987; Carcamo et al., 1989).

The promoter strength of RNA polymerase III transcribed genes has been shown to be affected by 5'-flanking sequences (reviewed in Geiduschek and Tocchini-Valentini, 1988). Transfer RNA genes usually show little flanking sequence homology either to other tRNA genes of the same family or to the flanking sequences of other families A comparison of three $tRNA^{Val}$ extragenic sequences revealed no significant homologies, indicating the separation of these genes early in vertebrate evolution (Arnold et al., 1986). It appears that the *in vitro* transcription of these tRNA genes strongly depends on extragenic control regions. As a consequence of different flanking regions, the transcription efficiencies vary by an order of magnitude among the genes for the major and minor valine tRNA genes and this reflects the concentrations of these two tRNAs *in vivo*. Two plasmids containing $tRNA^{Val}$ genes exhibit short stretches of homology in analogous positions upstream and downstream from the structural genes. It is unknown whether this is due to evolutionary relationship or functional constraint. Alu-type repeats were shown to occur about 500 and 800 bp downstream from the valine tRNA genes. The effect of flanking sequence on transcription was also demonstrated for the $tRNA^{Lys}$ gene family where two members of this human gene family with dissimilar flanking sequences were found to exhibit marked differences in

the efficiencies of expression (Doran et al., 1987). There are, however, examples of human tRNA genes which show an extremely high degree of flanking sequence homology. The 5'- and 3'-flanking regions of two independently isolated human-λ recombinants, each containing a single $tRNA^{Asn}$ gene, were shown to be over 90% homologous (Ma et al., 1984). Additionally, a comparison of two $tRNA^{Tyr}$ gene flanking sequences exhibited only 3 differences in over 350 bp of sequence information (van Tol et al., 1987; MacPherson and Roy, 1986). However, these $tRNA^{Tyr}$ genes may be alleles since the fragments show little flanking sequence homology to other known human $tRNA^{Tyr}$ gene sequences (MacPherson, 1988). The exact influence of flanking sequences on tRNA gene transcription for the majority of human tRNA genes is unknown.

1.3 The organization of the human genome

The haploid human genome is composed of 23 chromosomes (22 autosomal chromosomes and 1 sex chromosome, either X or Y) containing approximately $3.3x10^9$ base pairs comprising the primary nucleotide sequence. The morphology of these chromosomes is not common amongst the different chromosomes nor are the DNA sequences uniform or 'average' when compared to other regions of the same chromosome or to regions found on different chromosomes. Chromosomes are typically described as being divided into heterochromatic regions, referring to the regions of the centromeres and the telomeres as well as any secondary constrictions, and the euchromatic regions which are taken to be the rest of the genome. Differences in the function of various regions of the genome are caused by local differences in the primary nucleic acid sequence. Different regions of the genome may contain segments that are highly transcribed, highly supercoiled, highly methylated or are tightly bound to a variety of proteins required for the function of that section of the genome. Because of the division between nucleus and cytoplasm, the arrangements for gene expression

in eukaryotes must necessarily be different from those in prokaryotes. These differences are represented by differences in the organization of the eukaryotic genome as well as differences in the sets of effector molecules required to mediate efficient transcription of individual genes. Eukaryotic DNA is organized into structures called nucleosomes (Kornberg, 1974), comprised of approximately equal masses of DNA and histones. This structure functions partly to fold the vast expanse of DNA into a shape and size that allows packaging into the confines of the nucleus. Formation of stable complexes of transcription factors with promoters might affect the structure and assembly of chromatin (DNA plus bound histone molecules). Similarly, chromatin assembly might affect the outcome of competition for transcription factors among genes or families of genes. The importance of the RNA polymerase III transcription factors for the assembly of nucleosomes and the structure of chromatin on RNA polymerase III transcribed genes should be noted (reviewed in Wolfe and Brown, 1988). Morse (1989) has shown that RNA polymerase III, unlike RNA polymerase II, is unable to transcribe through a nucleosome, giving intragenic placement of TFIIIA and TFIIIC the key role in "reserving" a gene for transcription. The additional transcription factor TFIIIB may also act as an anchor to help retain TFIIIA and TFIIIC on the template while RNA polymerase III transcribes through their DNA binding domains (Brown, 1984).

Nucleic acid sequences found within eukaryotic DNAs can be loosely classified as (a) unique (approximately one copy per haploid genome), (b) moderately repetitive (~10-100,000 copies per haploid genome), or (c) highly repetitive (~1,000,000 copies per haploid genome - reviewed in Jelinek, 1982). The unique DNA fraction generally contains protein encoding sequences, although some protein coding genes are also represented more than once per haploid genome. Many other sequences, including human tRNA genes and pseudogenes (non-functional tRNA gene-like sequences) comprise the middle repetitive fraction of the human genome. Other families such as

the human Alu repeat sequence family (Rubin et al., 1980) contain as many as 500,000 members (Rinehart et al., 1981). The highly repetitive DNA fraction consists of clustered repetitions of relatively short sequence units (usually a few to a few hundred bp) that are thought not to be interspersed with other sequence types. They are generally considered to be structural components of chromosomes, residing mainly at centromeric and telomeric positions. These sequences may function to maintain the integrity of chromosome ends during DNA replication (Jager and Philippsen, 1989) and also to aid in the correct assortment of chromosomes during mitosis by serving as attachment sites for the mitotic spindle apparatus. This description of genomic organization of DNA sequences is not exhaustive and many other means of categorizing sequences exist, including the organization of DNA sequences as dispersed or clustered repeats. Other repeated sequences may show a variety of genomic organization patterns (e.g. human tRNA genes).

Genomic DNA can also be described in terms of it reassociation behavior in liquid hybridization experiments (e.g. $Cot_{1/2}$ analysis) or in solid phase hybridization experiments (Southern analysis). Repetitive sequences such as Alu or GT repeats that are found very frequently in the genome reanneal to their complementary sequences more quickly than those sequences that occur as single copies. Renaturation rate kinetics for the formation of duplex DNA provides only an approximate estimate of the number of members for each sequence family and provides little or no information on the genomic organization of different sequence families. Differences in functionally related sequences, due to mutational divergence, reduce the cross-renaturation rate as compared to the renaturation rate of perfectly complementary sequences (Bonner et al., 1973). Moreover, the renaturation rate of DNAs depends on the length of complementary sequences, which in the case of dispersed repeats also depends on the unknown details of the intervening sequences (Wetmur and Davidson, 1968). Renaturation rate studies of genomic DNA have been interpreted to mean that different

repetitive sequence families containing at least 10 to 10,000 members are present in a variety of eukaryotic DNAs.

1.4 The genomic organization of human tRNA genes

By using saturation hybridization experiments, Hatlen and Attardi (1971) estimated that there are approximately 1300 tRNA genes in the haploid human genome. These are thought to encode 60-90 different species of tRNA (Lin and Agris, 1980). This suggests an approximate average of 65 copies for each isoaccepting tRNA gene family or approximately 10-20 copies for each chromatographically distinct tRNA. Estimation of gene numbers for human tRNA genes varies for different tRNA gene families. Comparison of relative hybridization intensities for hybridization of $tRNA^{Asn}$ specific probes to copy number standards suggested that there are at least 60 $tRNA^{Asn}$ genes (Buckland, 1989). Valine tRNA genes have been detected on 13 different fragments in *Eco*RI digests (Arnold et al., 1986) while $tRNA_i^{Met}$ genes have been detected on 12 fragments (Santos and Zasloff, 1981). Similarly, van Tol and Beier (1988) observed 12 different $tRNA^{Tyr}$ containing fragments in *Eco*RI digests of human genomic DNA. The importance of hybridization conditions for reasonable estimates of copy number was demonstrated by Scheller et al (1981). The apparent copy number of one cloned repeat sequence varied with the stringency of hybridization. Under stringent conditions, there were an estimated 20 copies per haploid genome, whereas under non-stringent conditions the apparent copy number exceeded 400.

In humans, tRNA genes can be found as members of complex multigene clusters containing different isoaccepting tRNA genes (Roy et al., 1982; Pirtle et al., 1986; Craig et al., 1989) or they may occur as clusters containing similar tRNA genes. MacPherson and Roy (1986) demonstrated the first example of this latter arrangement where a λ recombinant was shown to contain at least four $tRNA^{Tyr}$ genes. Another λ recombinant was described that contained two $tRNA^{Phe}$ genes arranged as tandem

repeats separated by ~9 kb and two $tRNA^{Lys}$ genes arranged as inverted repeats separated by ~5 kb within a 13.8 kb fragment (Doran et al., 1987). Chang et al. (1986) isolated a lambda clone containing a 20 kb human DNA segment that was found to harbor a cluster of four tRNA genes. The genes were found to be arranged as two tandem pairs, separated by 3 kb. A proline tRNA gene is separated from a leucine tRNA gene by a 724 bp intergenic region in the first pair, and a second proline tRNA gene is 316 bp from a threonine tRNA gene in the second pair, with the leucine tRNA gene being of opposite polarity to the other three genes. An Alu-like element was found to occur within a 2.0 kb DNA fragment, approximately 0.7 kb from the tRNA gene cluster. No apparent homologies occur between the 5'-flanking sequences of these genes. The arrangement in clusters is most likely not representative for the genomic organization of the majority of mammalian tRNA genes, since clusters can be detected more easily than solitary genes, if as is usually the case, unfractionated tRNA or tDNA (a DNA sequence complementary to the tRNA sequence) containing a cluster is used as a hybridization probe to screen λ libraries. Single tRNA genes of the same or different acceptor specificity can usually be found at other locations in the genome. Accordingly, some examples of human tRNA genes apparently occurring as individual genes have also been described, including genes for $tRNA^{Gly}$ (Doran et al., 1988; Shortridge et al., 1985; Pirtle et al., 1986), $tRNA^{Tyr}$ (van Tol et al., 1987), and $tRNA^{Asn}$ (Ma et al., 1984). Recently, Morrison et al., (1991) demonstrated that one of the single $tRNA^{Gly}$ genes described by Doran et al., (1988) was linked to the tRNA gene heterocluster previously described by Roy et al., (1982)

Chromosomal mapping of many human DNA fragments has been performed by hybridization analysis of mouse-human somatic cell hybrids. The genotypic composition of these hybrid cells is very unstable. The mouse chromosome complement is usually retained while the human chromosomes are lost randomly. By performing cytogenetic and enzymatic analysis of hybrid clones it is possible to

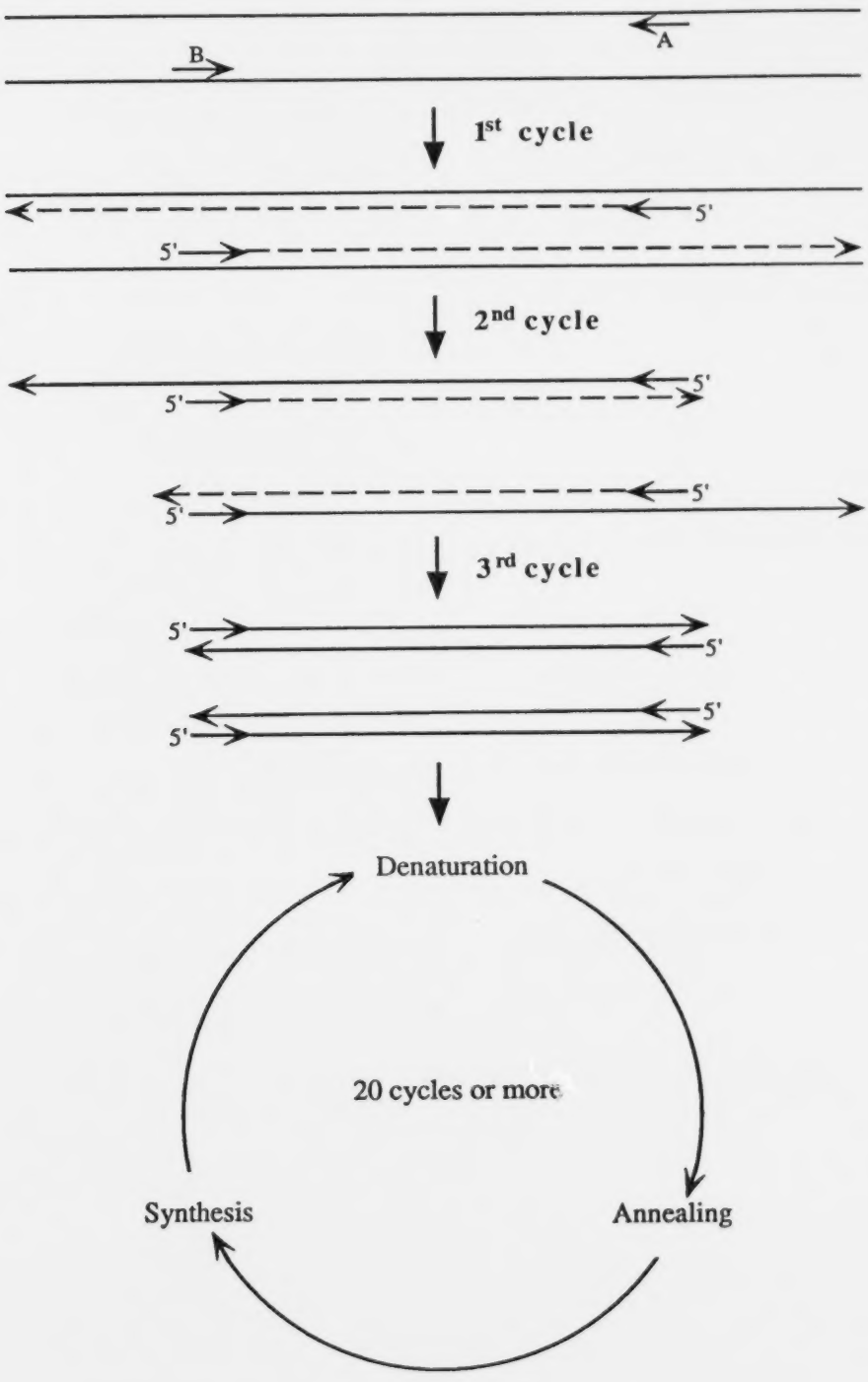
determine which human chromosomes have been retained. DNA isolated from somatic cell hybrids which possess different assortments of human chromosomes is used in the construction of chromosomal mapping panels. The pattern observed for the hybridization of human specific probes to the mapping panels therefore allows the assignment of the probe DNA to a particular human chromosome. The utility of using the unique flanking sequences surrounding human tRNA genes as hybridization probes for determining chromosomal location was first demonstrated by Naylor et al. (1983). Two nonallelic human $tRNA_i^{Met}$ genes were assigned to chromosome 6 by filter hybridization of DNA from human-rodent somatic cell hybrids. Both genes were assigned to the p23-q12 region. These results raised the possibility that other $tRNA^{Met}$ genes may be syntenic with these two tRNA genes. Two human DNA clones containing asparagine tRNA genes were used to determine the genomic organization of $tRNA^{Asn}$ multigene family in humans. One of the clones used in this study also contained a gene for U1 RNA, so the organization of the two multigene families was compared. The majority of the human $tRNA^{Asn}$ genes were shown to map to the same chromosome bands, as do the U1 RNA true genes and class I pseudogenes located on the short and long arms, respectively, of chromosome 1. The chromosomal locations of these two gene clusters were determined using a somatic-cell hybrid minipanel. The use of repeat-unit DNA polymorphisms showed that one tRNA gene clone maps to the short-arm gene cluster and the other to the long-arm gene cluster. The short- and long-arm gene clusters were shown to be structurally related. Because of these results it was suggested that the split into two distinct loci was facilitated by a pericentric chromosome inversion. Pulsed field gel electrophoresis revealed that fragments of molecular weight of 180 kb were common to each unit (or multiples of units). It has been proposed that these two gene families are closely linked on repeat units (or multiples of units) of 180 kb in size (Buckland, 1989).

McBride et al. (1989) determined the chromosomal locations of three cloned human DNA fragments, encompassing tRNA genes, by Southern analysis of human-rodent somatic cell hybrid DNAs with subfragments from these cloned genes and their flanking sequences used as hybridization probes. These three DNA segments have been assigned to three different chromosomes. A DNA segment containing a single $tRNA^{Gly}$ gene was shown to be located on human chromosomes 1, but it also segregated with a clone containing two $tRNA^{Gly}$ genes that hybridized to DNA from chromosome 16. Homologous sequences to these DNA fragments were postulated to be located on chromosome 14 and on a separate locus on chromosome 1. The third DNA fragment, containing single copies of $tRNA^{Thr}$, $tRNA^{Pro}$ and $tRNA^{Val}$ genes, was determined to be located on chromosome 5. By using *in situ* hybridization, a tRNA gene heterocluster containing genes for $tRNA^{Arg}$, $tRNA^{Gly}$, $tRNA^{Lys}$, $tRNA^{Leu}$ and $tRNA^{Gln}$ have been shown to be located on chromosome 17 (Morrison et al., 1991). Other studies have shown that an "opal suppressor" tRNA gene (later described as phosphoserine tRNA or selenocysteine tRNA) and pseudogene are located on chromosome 19 and 22, respectively (McBride et al., 1987). Glutamate tRNA genes have been assigned to two loci on chromosome 1p36 (Boyd et al., 1989). These studies, combined with previous results, indicate that tRNA genes and pseudogenes are dispersed on at least eight different human chromosomes and suggest that these sequences will probably be found on most, if not all, human chromosomes. Because of the short (8-12 nucleotide) direct terminal repeats observed to be flanking many of the dispersed tRNA genes it was suggested that a significant proportion of these genes have arisen by an RNA-mediated retroposition mechanism (McBride et al., 1989).

1.5 The polymerase chain reaction

In order to develop a more comprehensive understanding of the structure, location, and transcriptional regulation of human tRNA genes, it is necessary to acquire more examples of genes for each tRNA family. Only a few examples are known for most of the estimated 10-20 tRNA genes for each isoaccepting tRNA gene family. For a few gene families, such as $tRNA^{Cys}$, $tRNA^{His}$, and $tRNA^{Trp}$, no examples of genes for these tRNAs have been described. The Polymerase Chain Reaction (PCR) developed by Mullis and Faloona (1987; Saiki et al., 1988), represents a very powerful tool for the analysis of nucleic acids. This thesis describes various strategies, based on PCR, which have been adapted to allow the accelerated isolation of novel tRNA genes and flanking regions for further study. The fragments so isolated can be used directly for study, or can be used as specific probes to isolate larger fragments from human genomic libraries. PCR, as originally described, involves the selective amplification *in vitro* of sequences up to ~4 kb or more in length, by sequential rounds of template denaturation, primer annealing and extension from the annealed primers using a DNA polymerase. The primers used in this reaction are positioned in the flanking regions surrounding a region of interest and are oriented in opposite and overlapping directions (see Figure 3). A DNA polymerase, usually the thermostable DNA polymerase (*Taq*) isolated from *Thermus aquaticus*, then recognizes this as a suitable substrate for primer extension in the direction of the second primer site and eventually the second primer binding site is synthesized. This method thus allows the exponential amplification of DNA between the two primer binding sites because both strands of a given DNA segment are used as templates. This reaction is simple, fast and extremely sensitive as shown by the fact that the DNA (or RNA) content of a single cell is often suitable for the detection of a specific sequence. The usual prerequisite for the use of PCR is *a priori* knowledge of sequences surrounding the fragment of DNA that one wishes to amplify in order to design suitable primers.

Figure 3 - Schematic representation of the Polymerase Chain Reaction. Primers A and B that are oriented in overlapping directions are annealed to the denatured template DNA in the first round. Template directed enzymatic extension from these primers is then carried out, usually by a thermostable DNA polymerase. The synthesized products are extended to include the other primer binding site, so that the products of synthesis in each round are suitable targets for the binding of the opposite oligonucleotides in the subsequent round of PCR. Each round of PCR thus consists of template denaturation at 94°C to 96°C, usually for 15 to 60 seconds, followed by primer annealing to the template at a temperature suitable for a given pair of primers, usually for 15 to 60 seconds. A cycle of PCR is then completed by the extension from the primers at 70°C to 74°C for 30 seconds to 5 minutes. Each cycle of PCR (denaturation/primer annealing/primer extension) is repeated 20 to 35 times. By this means the predominant product of amplification is the region of DNA that is found between the two primer binding sites. Since the product of primer extension of each round is a suitable target for synthesis in the next round, the exponential amplification of specific DNA sequences is possible (Adapted from Vosberg, 1989).


A
B
1st cycle
5'
5'
2nd cycle
5'
5'
5'
5'
3rd cycle
5'
5'
5'
5'
Denaturation
20 cycles or more
Synthesis
Annealing

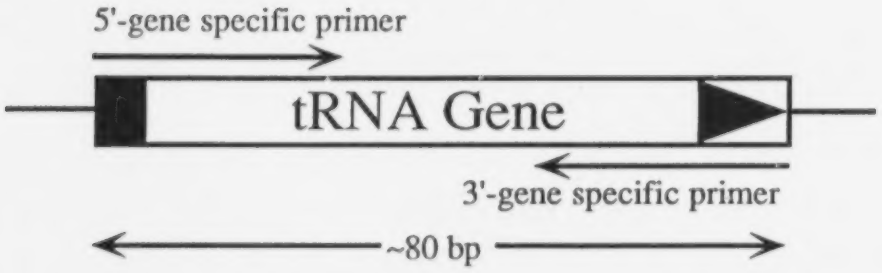
1.5.1 tRNA PCR

Transfer RNA genes represent a favorable system for the recovery of novel human genomic sequences because the tRNA coding sequence is usually highly conserved for most isoaccepting tRNA gene families. Moreover, because the flanking sequences surrounding these genes are generally quite diverse, these gene sequences could be very useful for providing unique genomic mapping sites. As previously described, human tRNA genes may be found as clusters of isoaccepting tRNA genes, as was shown for the human $tRNA^{Tyr}$ gene family (MacPherson and Roy, 1986). This arrangement of genes allows the use of "tRNA PCR" where PCR is performed using tRNA gene-specific primers which are directed at the 5'- and/or 3'-conserved gene sequence(s), and are oriented through the gene (see Figure 4). PCR is done using both primers to recover the tRNA gene sequence alone (and intron sequence, if present) and also to isolate tRNA genes arranged as tandem repeats. Alternatively, by performing PCR using only the 5'- or the 3'-primers individually in a single primer PCR, it should be possible to amplify tRNA sequences that occur as inverted repeats relative to each other. Amplification products due to other arrangements of genomic sequences, including tRNA pseudogenes or repeat sequences, should also be possible. Similarities between tRNA genes and repeat sequences have been observed. Sakamoto and Okada (1985) demonstrated close structural resemblances between several mammalian highly or moderately repetitive sequence families and bovine tRNA genes coding for $tRNA^{Lys}$, $tRNA^{Phe}$ and $tRNA^{Gly}$. This is supported by the close sequence similarities observed between the $tRNA_i^{Met}$ gene sequence and the Alu sequence family, and between the mouse B2 families (Krayev et al., 1982) and the genes coding for $tRNA^{Ala}$ and $tRNA^{Ser}$ in the prosimian *Galago crassicaudatus* (Daniels and Deininger, 1985). The nature of amplification products from tRNA PCR may reveal a glimpse into the evolutionary history of human tRNA genes, and may lend support for previously proposed mechanisms of tRNA gene duplication and dispersal in the genome.

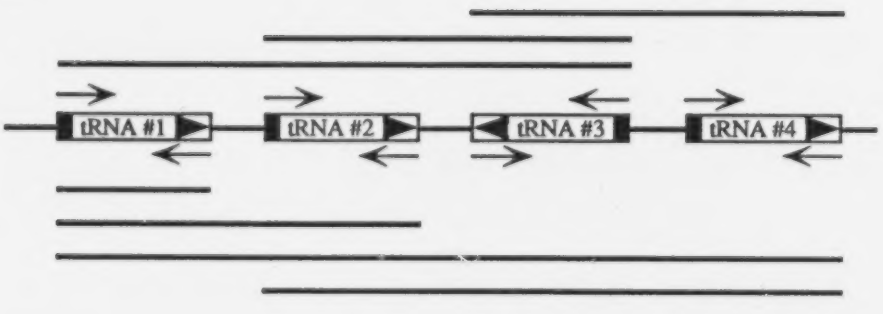
Figure 4 - Schematic representation of tRNA PCR.

Figure 4(A) shows the position and relative orientation of two tRNA gene specific primers that are used for PCR. Each primer starts at the 5' or 3' terminus of the tRNA gene coding sequence and extends into the gene for 21 to 25 nt. The region that would be expected to contain introns (usually one nt 3' to the anticodon) is not contained in either of the primer sequences. Figure 4(B) shows possible arrangements of tRNA genes relative to one another in a hypothetical tRNA gene cluster contained in the genome. The arrows represent the tRNA gene specific primers. Solid lines indicate possible amplification products that would arise from this arrangement of tRNA genes for each combination of primers. When only the 5' gene-specific primer is used in a PCR, amplification is possible between two tRNA genes that are arranged as inverted repeats. Similarly, amplification between inverted repeats of the opposite orientation is possible when the 3' gene-specific primer is used alone in a PCR. When both primers are included in a PCR, amplification of tRNA gene sequences that are found arranged as tandem repeats is possible. In this latter scenario, amplification of the tRNA gene sequence only (and intron sequence, if present) is expected to be the most prevalent amplification event.

A


5'-gene specific primer
tRNA Gene
3'-gene specific primer
~80 bp

B


tRNA #1
tRNA #2
tRNA #3
tRNA #4

1.5.2 Alu-tRNA PCR

It is possible that neighboring Alu-type repeats are a general feature of human tRNA gene loci. Alu sequences represent the major family of short interspersed repeat DNAs (SINEs). Additionally, like tRNA genes, Alu sequences contain conserved promoter sequences for RNA polymerase III. The estimated distance between Alu sequences (~4 kb; Britten et al., 1988) varies considerably as evidenced by the preponderance of these sequences in certain chromosomal regions. Alu enrichment is known to occur in the vicinity of many tRNA genes. Analysis of a human gene cluster coding for $tRNA^{Phe}$ and $tRNA^{Lys}$ revealed eight regions of Alu homology on a DNA fragment of 13.8 kb (Doran et al., 1987). Similarly, an 18.5 kb DNA segment containing a cluster of three tRNA genes was also shown to harbor at least nine Alu family members (Craig et al., 1989). Remarkably, in one human tRNA gene heterocluster an Alu element overlapping a valine tRNA gene was demonstrated (Shortridge et al., 1989). The role, if any, of Alu sequences in the modulation of tRNA gene expression is unknown.

Nelson et al., (1989) first described the use of the genomic arrangement of Alu sequences for the recovery of specific human sequences using the PCR technique. This was done for complex DNA sources such as mouse/human somatic cell hybrids or Yeast Artificial Chromosomes (Burke et al., 1987). This technique, termed "Alu PCR", utilizes primers that are directed at conserved regions within the Alu sequence. The success of this method depends upon the serendipitous arrangement of two Alu repeats flanking a region of interest that one wishes to amplify. While this method has worked well for some DNA sources it failed to provide appreciable amplification of specific sequences directly from human genomic DNA. The limitations of this technique include the requirement for two sequences to be orientated in the correct manner flanking a sequence of interest within an amplifiable distance. An equally

severe limitation was the significant "amplification noise" that was due to Alu-Alu amplification from other regions of the genome.

"Alu-tRNA PCR" was designed to overcome this lack of specificity and low signal to noise ratio. In this technique, one primer is directed at a conserved region of an Alu repeat element and another is directed at a conserved region within a tRNA gene (see Figure 5). The problem of two sequences flanking a region of interest is now limited to the probability of a single occurrence. The choice of multiple primers in opposite orientations increases the likelihood of amplification of a desired sequence. Additionally, if the tRNA gene-specific primer is biotinylated it will allow selection of tRNA gene-specific amplification products from among the predominant Alu-Alu products which constitute background noise.

1.5.3 Inverted PCR

The limitation of PCR for the recovery of novel sequences from the human genome is that it is highly dependent on knowledge of two suitable primer binding sites before the experiment is performed. "Inverted PCR" (Ochman et al., 1988; Triglia et al., 1988) is a procedure that allows the amplification of sequences that lie outside the boundaries of known sequences. This approach depends on the inversion of the sequence of interest by circularization and re-opening at a different site. Genomic DNA is first cleaved with a restriction endonuclease and the fragments ligated under conditions that favor circularization. PCR is performed using primers that are arranged in an opposite orientation to that used for normal PCR. The advantage of this PCR method is that it allows amplification of novel sequences when very little sequence information is known. Because the coding regions of tRNA genes are highly conserved, this gene family lends itself well to the use of inverted PCR for the recovery of novel sequences flanking the tRNA sequences. A schematic representation of this method, as applied to the recovery of tRNA genes, is shown in Figure 6.

Figure 5 - Schematic representation of Alu-tRNA PCR.

Figure 5(A) shows the possible arrangements of Alu repeat sequences relative to a tRNA gene. Small arrows represent PCR primers that are specific for the Alu repeat sequence or for the 5' or 3' regions of a tRNA gene. Solid lines indicate possible amplification products for PCR performed using these primers. Amplification between the 5' gene-specific primer and two Alu repeat sequences is shown in (1) and (2). Amplification between the 3' gene-specific primer and two Alu repeat sequences is shown in (3) and (4). The most prevalent amplification is expected to occur between Alu repeat sequences as shown in (5) and (6). To overcome the extremely high background amplification between Alu sequences, a two stage PCR was performed using a biotinylated tRNA gene-specific primer in the first stage (figure 5(B)). Those products that are the result of amplification between a tRNA gene sequence and a Alu sequence should be retained by selection using streptavidin-agarose or other solid phase streptavidin derivatives. These selected sequences should be suitable templates for PCR using combinations of primers similar to those used in the PCR that generated each fragment. If non-biotinylated primers are used in the second stage PCR, cloning of amplification products should be possible.

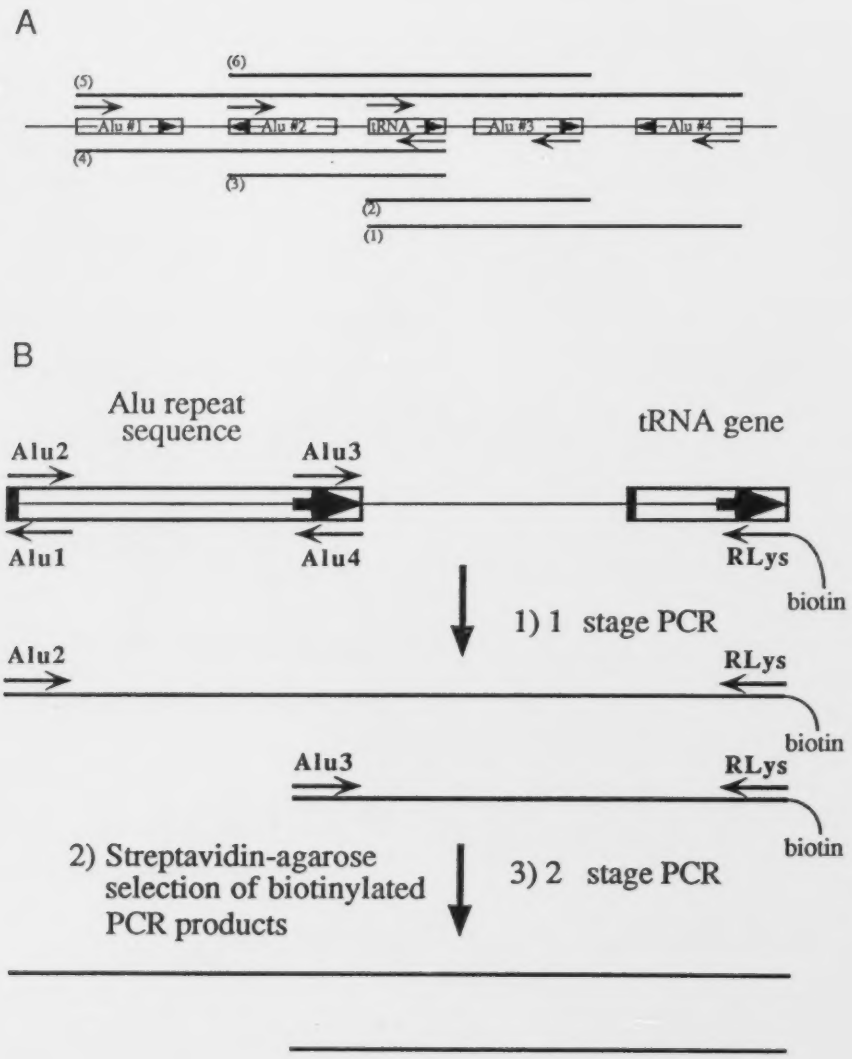
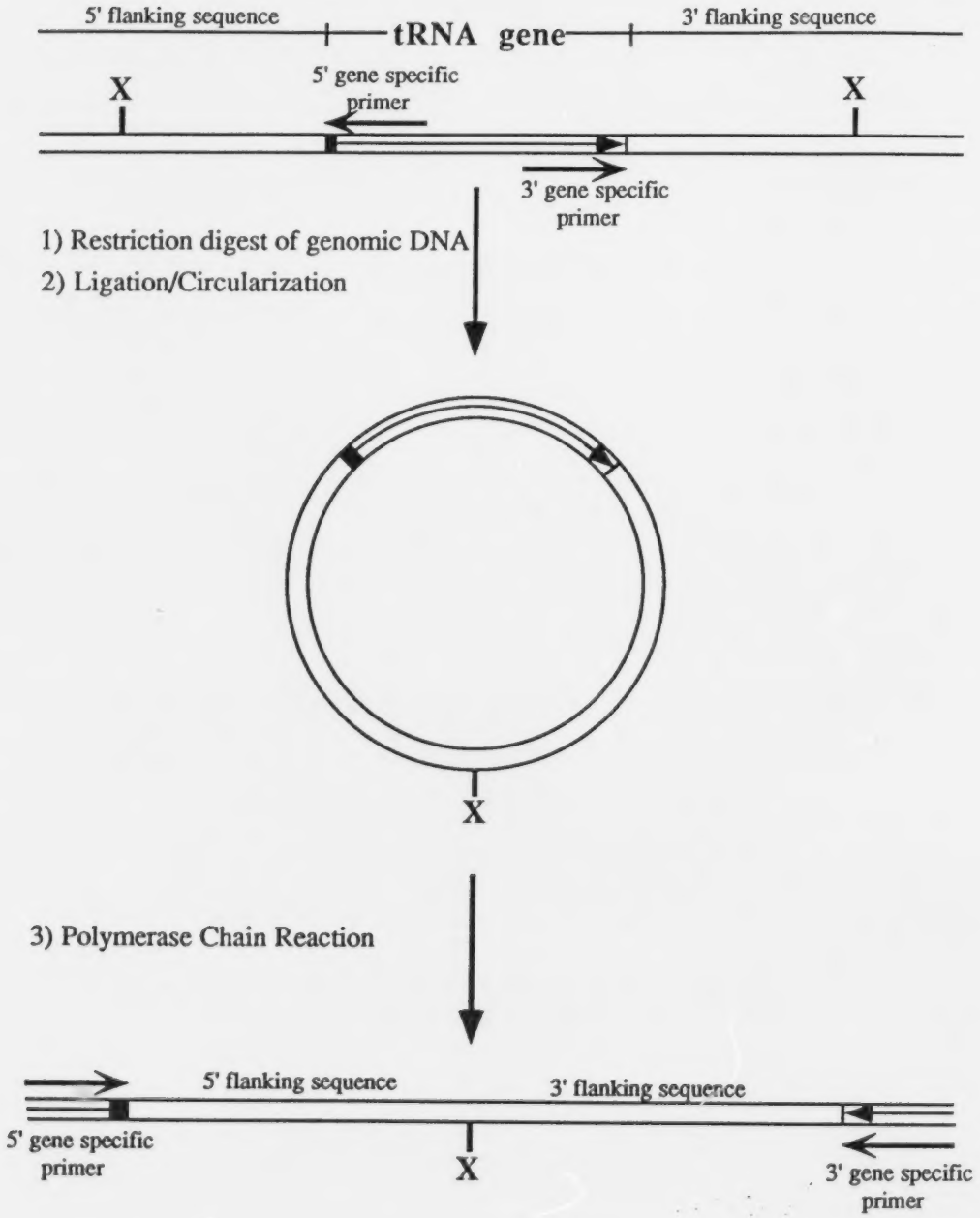

A
(6)
(5)
Alu #1
Alu #2
tRNA
Alu #3
Alu #4
(4)
(3)
(2)
(1)
B
Alu repeat sequence
tRNA gene
Alu2
Alu3
Alu1
Alu4
RLys
biotin
1) 1 stage PCR
Alu2
RLys
biotin
Alu3
RLys
biotin
2) Streptavidin-agarose selection of biotinylated PCR products
3) 2 stage PCR

Figure 6 - Schematic representation of Inverted PCR.

This diagram shows a representation of a single tRNA gene, plus its flanking sequence as it exists in the genome. The DNA is digested with a restriction endonuclease (X) that cleaves on both sides of the tRNA gene within a amplifiable distance (~4 kb). The digested DNA is then ligated under conditions that favor the circularization of the DNA. In this manner, two tRNA gene specific primers (arrows) that were oriented in non-overlapping directions on the genome are now arranged in overlapping directions on the circularized template. PCR is then performed. The amplification products of such a PCR should contain both the 5' and 3' coding sequences for a tRNA gene without the coding sequence that was contained between the primer binding sites in the tRNA gene. In addition, various amount of both the 5' and 3' flanking sequence should be found on a contiguous fragment, delineated by the recognition site for the enzyme that was used in the restriction endonuclease digestion step.


5' flanking sequence
tRNA gene
3' flanking sequence
X
5' gene specific primer
X
3' gene specific primer
1) Restriction digest of genomic DNA
2) Ligation/Circularization
X
3) Polymerase Chain Reaction
5' flanking sequence
3' flanking sequence
5' gene specific primer
X
3' gene specific primer

The major difficulty with inverted PCR is the lack of control over the circularization step. The theoretical basis of ligation of a linear polymeric molecule in solution (Jacobson and Stockmayer, 1950) has been experimentally tested for bacteriophage λ DNA (Wang and Davidson, 1966) and smaller DNA fragments (Dugaiczyk et al., 1975). This theoretical treatment results in the calculation of a parameter, j, that is the effective concentration of one end of a long polymer in the neighborhood of the other end. This parameter is given by the equation:

$$j = (3/2\pi lb)^{3/2} \text{ ends per ml,}$$

where b is length of a series of rigid segments joined by freely movable joints, comprising a polymer of length l.

The value for b (for double stranded DNA) has been estimated to be 7.2×10^{-2} μm (Hearst and Stockmayer, 1962). This leads to the simplification :

$$j = 63.4/(kb)^{1/2} \text{ μg/ml,}$$

where kb is the length of a particular DNA fragment in kilobase pairs. If ligation is carried out at a DNA concentration, i, which is less than j, formation of circles will be favored. If i > j, multimer formation will be favored. At a given concentration, i, the fraction of circles formed is predicted by the equation:

$$\% \text{ circles} = j/j+i \times 100.$$

Experimental evaluation of these estimates has been performed by Collins and Weissman (1984). Their analysis of circularization efficiency by electrophoresis, however, showed that a significant degree of circularization occurs even at relatively high DNA concentrations (30 μg/ml).

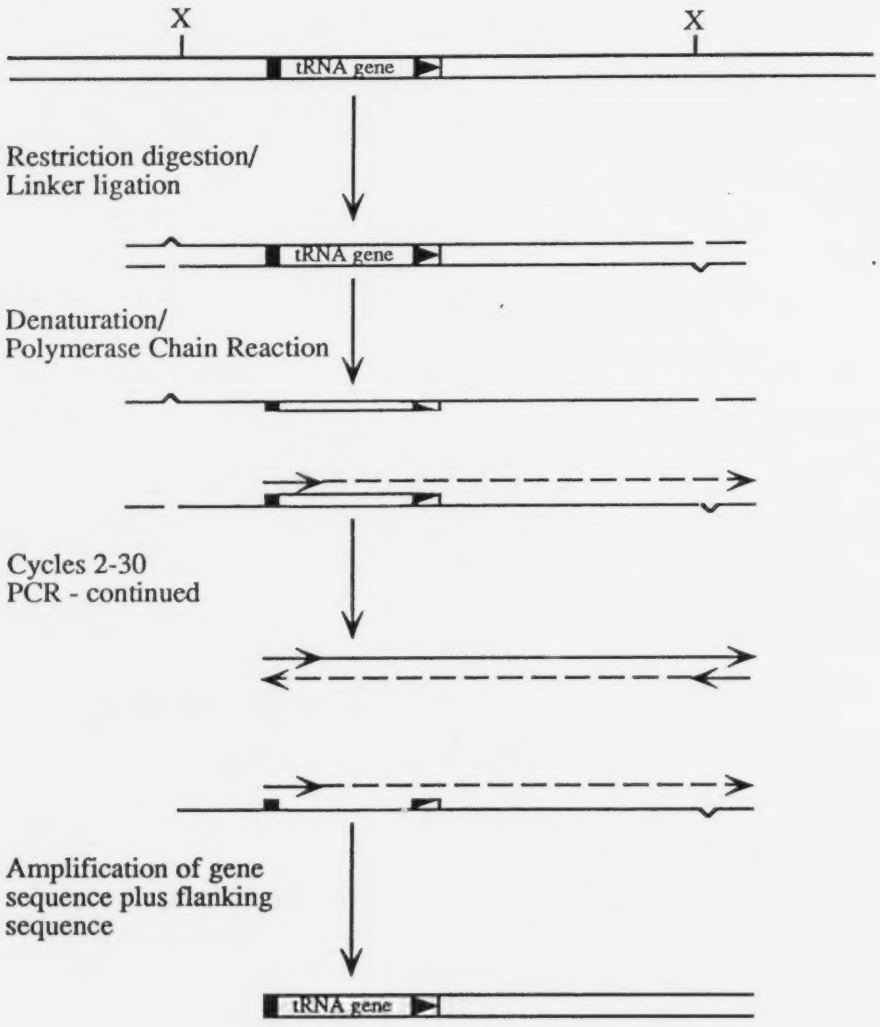
1.5.4 Linker PCR

Another method that may prove useful for the isolation of novel tRNA gene-containing sequences is a method termed "Linker PCR". In this technique, total genomic DNA is digested with a restriction endonuclease that generates blunt ends. A second primer binding site suitable for PCR is then ligated to the ends of the DNA fragments in the form of double stranded linker oligonucleotides. A PCR is then performed using a tRNA gene-specific primer and a second primer which is complementary to one of the two strands of the linker. In order to prevent the high background amplification which would be expected to occur between linkers ligated to both ends of the digested DNA molecules, the PCR is done in two stages. The first stage PCR utilizes a biotinylated gene-specific primer so that the desired amplification products can be separated from the background noise by subsequent selection with streptavidin-agarose or other solid phase streptavidin derivatives. A second stage PCR is then performed to amplify the DNA fragments that contain the gene-specific primer binding site. The advantage to this method is that it is possible to amplify specific sequences with only limited sequence information and without the constraints of requiring adjacent tRNA genes, Alu sequences, or pairs of restriction endonuclease sites. Because human tRNA gene coding regions are very conserved while the flanking sequences usually are quite diverse, linker PCR involving human tRNA gene sequences may thus provide a useful system for recovering novel unique sequences directly from the genome. By using six primers (two biotinylated gene-specific primers, two non-biotinylated gene-specific primers and two linker primers) it should be possible to recover numerous sequences in both the 5'- and 3'- flanking directions from each member of a given tRNA gene family (see Figure 7).

"Ligation Meditated PCR" was first described by Mueller and Wold (1989) as a means of performing *in vivo* footprinting of muscle specific enhancers in humans. Pfeifer et al. (1989) were able to use this technique to generate sequence ladders using

Figure 7 - Schematic representation of the Linker PCR strategy.

In the first step, total genomic DNA is digested with a restriction enzyme. Linkers are ligated to the ends of the digested DNA. Note that because the primers are not 5'-phosphorylated, concatenation of the linkers is not possible because ligation only occurs between the 3'-OH of the linkers and the 5'-phosphate of the digested DNA. Ligation of multiple genomic fragments should be detectable by the presence of the restriction recognition sequence for the enzyme that was used in the original digestion step. The first round of PCR involves extension of the gene-specific primer to create the annealing site for the second primer, one of the linker primers. Subsequent rounds of PCR allow amplification of the tRNA gene plus either the 5' flanking sequence or the 3' flanking sequence depending upon the orientation of the gene specific primer. The use of primers that are specific for the flanking sequence regions should allow 'walking' from this gene locus by this method.


X
X
tRNA gene
Restriction digestion/
Linker ligation
tRNA gene
Denaturation/
Polymerase Chain Reaction
Cycles 2-30
PCR - continued
Amplification of gene
sequence plus flanking
sequence
tRNA gene

1 µg of uncloned human DNA. This method was subsequently extended to recover a promoter sequence from shark genomic DNA by using a linker primer and nested gene-specific primers in multiple rounds of PCR. The staggered primers were necessary in order to generate the required specificity for recovery of this unique sequence from the genome (Fors et al., 1990). A similar technique termed "oligo-cassette mediated PCR" utilized biotin selection of amplification products as a means for genomic walking (Rosenthal and Jones, 1990). The use of linkers for PCR has also been described for the recovery of cDNA sequences directly from mRNA using a poly (dT) tail oligo that is complementary to the poly (dA) tail seen on eukaryotic mRNA (Loh et al., 1989).

1.6 Objectives of this study

One objective of this study was to determine the chromosomal locations of some human tRNA gene-containing DNA fragments that had been previously characterized in this laboratory using standard methods (Doran et al., 1987; Doran et al., 1988; Roy et al., 1982; MacPherson and Roy, 1986). The assignment of tRNA genes to specific chromosomal regions may be important in the continuing accumulation of physical map markers for the human genome. This is significant for two reasons:

i) Genetic linkage mappers are trying to identify markers at closely spaced intervals throughout the human genome, so that disease genes can be quickly localized to a fairly narrow region.

ii) The proposed effort to sequence the entire human genome requires the construction of a set of markers that completely spans the human genome. The ability to locate individual genes of multigene families such as tRNA genes may aid in this process.

In this study, various probe preparation methods and hybridization conditions were analyzed to determine the best procedure for obtaining hybridization to specific

sequences within human genomic DNA. To determine the chromosomal locations of human tRNA gene-containing fragments, hybridization analysis of mouse/human somatic cell hybrid chromosomal mapping panels was done for four human-λ recombinants.

A second objective was to develop new methods for the isolation of human tRNA genes and other functional sequences directly from the human genome. The significance of genomic patterns of human tRNA gene dispersion and organization in regard to tRNA gene expression is unknown. Since eukaryotic tRNA genes exist as multigene families of independently transcribed genes, it is thus important to determine the organization and structure of members of human tRNA gene families and to determine whether the location or arrangement of these genes is significant in terms of the regulation of tDNA expression. Moreover, the nature of the organization of tRNA gene families may allow the use of molecular techniques to visualize individual specific, DNA fingerprints (Jeffreys et al., 1985). This concept of DNA fingerprints has already found practical usage in immigration investigations, paternity identifications, as well as in forensic applications. Additionally, the product of a constructed human suppressor $tRNA^{Lys}$ gene is capable of suppression of an amber mutation encoded in β^{o}-thalessemia mRNA without interfering noticeably with normal host-cell mechanisms for translation termination (Temple et al., 1982). Reasons for this effect or lack thereof on host cell translation may have some basis in the regulation of tRNA transcription. Further studies along these lines could be greatly facilitated if more examples of tRNA genes for each gene family were known.

Traditional methods of tRNA gene isolation, such as characterization of human-λ recombinant libraries, are very labor intensive, especially when examining many tRNA gene loci. To date, only a few examples of tRNA genes for most gene families are known. The use of PCR allows the accelerated isolation and description of tRNA gene sequences. Four strategies, based on PCR, were utilized to amplify tRNA genes

or pseudogenes. Two of these techniques, "tRNA PCR" and "Inverted PCR", were applied successfully to the recovery of sequences for the $tRNA^{Tyr}$ and $tRNA^{Phe}$ gene families. These techniques, along with "Alu-tRNA PCR" and "Linker PCR", can be expanded to examine genes for all isoaccepting tRNA gene families for which known sequences are available. By using these methods, it should be possible to isolate all members of a tRNA gene family for each isoaccepting tRNA species. Ultimately it is hoped that the characterization of novel sequences flanking human tRNA genes will be useful for transcriptional regulation and genome mapping studies. The techniques described and/or developed herein are widely applicable to the study of other DNA sequences of varying function and to DNA from a variety of species.

2. METHODS AND MATERIALS

2.1 Laboratory materials and enzymes

Restriction enzymes were purchased from Boehringer Mannheim, New England Biolabs or Bethesda Research Laboratories and were used in the appropriate buffer as supplied by the manufacturer. Restriction endonuclease digests were done using ~2-10 units of enzyme per μg DNA, at the recommended temperature for ~3 hr. T4 DNA ligase was purchased from Bethesda Research Laboratories or from Boehringer Mannheim. T4 polynucleotide kinase was purchased from Pharmacia. Streptavidin-agarose was purchased from Sigma. Taq DNA polymerase was from Boehringer Mannheim. Nucleotides used for sequencing and for PCR were purchased from Pharmacia in the lyophilized form and subsequently reconstituted to 10 mM as a stock solution. The radioisotopically labelled compounds [α^{32}-P]dATP and [γ^{32}-P]-ATP were from New England Nuclear Ltd or ICN Biochemicals Inc. X-ray film was Kodak X-omat type AR film. PCRs were performed using the Techne PHC-2 thermocycler. All oligonucleotides were synthesized on a Applied Biosystems model 381A or 391EP DNA synthesizer at the Department of Microbiology (University of Alberta) DNA synthesis facility. Agarose and acrylamide were purchased from Boehringer Mannheim.

2.2 Isolation of human and mouse genomic DNA

Total human genomic DNA was isolated from buffy coat preparations of human blood obtained from the Alberta Red Cross. High molecular weight DNA was released by SDS lysis (final concentration 0.5%) of human leukocytes. Lysates were extracted twice with an equal volume of phenol, once with phenol:chloroform (1:1), and once with an equal volume of chloroform. To the aqueous phase was added one tenth volume of 3 M NaOAc. Two volumes of 95% ethanol were carefully layered onto the

aqueous phase. Large strands of genomic DNA were then carefully spooled from the aqueous phase into the alcohol phase using a hooked, silanized Pasteur pipette. DNA was then redissolved to a final concentration of 0.5-2.0 μg/μl. DNA size was confirmed by field inversion gel electrophoresis (Carle et al., 1986).

Mouse genomic DNA was isolated from the livers of Balb/c and C3H mice. Livers were removed and immediately transferred to liquid nitrogen in a mortar and ground to a fine powder using a pestle. The pulverized livers were resuspended in lysis buffer (1% SDS; 10 mM Tris-HCl/pH 8.0; 1 mM EDTA) and were subsequently extracted twice with phenol, once with phenol/chloroform (1:1 ratio), and finally once with chloroform. DNA was precipitated with ethanol and isolated as previously described for the isolation of human genomic DNA.

2.3 Preparation of DNA probes

DNA probes used for hybridization to human genomic DNA were prepared by either the M13 single stranded hybridization probe procedure of Hu and Messing (1982) or were labelled by the random primer method as described by Feinberg and Vogelstein (1983, 1984). Oligonucleotide probes were labelled by T4 polynucleotide kinase and [γ-^{32}P]ATP as described by Maxam and Gilbert (1980). All probes were purified by passage through a Sephadex G-50 column. Estimation of the specific activity of radioactively labelled probes was done by checking a measured aliquot of the purified probe preparation by liquid scintillation spectrometry for Cerenkov radiation.

2.4 Human tRNA gene localization studies

2.4.1 Standardization of hybridization conditions

DNA for hybridization experiments was transferred by the method of Southern (1975) to nylon membrane using the modifications described by Rigaud et al., (1987). The conditions for hybridization of tRNA gene-containing probes prepared by 5'-end labeling, M13 primer extension and random primer probe preparation methods were investigated. Unless otherwise noted all hybridizations were performed in the same solution as used for the prehybridization step. Prehybridizations were performed for 6-16 hr at 42-65°C in standard hybridization solution (5x Denhardt's solution/6x SSPE/0.1% SDS - 1x Denhardt's is 0.5 mg/ml bovine serum albumin, 0.5 mg/ml ficoll 400, 0.5 mg/ml polyvinylpyrrolidone; 1x SSPE is 150 mM NaCl, 10 mM NaH_2PO_4, 1 mM Na_2EDTA, pH 7.6) plus ~100 µg/ml sonicated *E. coli* DNA.

A. Hybridization of oligonucleotides to human genomic DNA: Approximately 5 ng of Tyr1 DNA (ATAGCTCAG[C]TGGTAGAGC) or Tyr2 DNA (ATAGCTCAG[T]TGGTAGAGC) was labelled as previously described and hybridized at 42°C to 10 µg human genomic DNA digested with either *Bam*HI, *Bgl*II, *Eco*RI, or *Hin*dIII. Membranes were washed at 45°C in 2x SSPE/0.1% SDS and exposed to X-ray film for 5 days at -80°C using 2 Lightning Plus intensifying screens (Dupont). Tyr1 and Tyr2 are 19 nucleotides in length beginning at position 7 of the $tRNA^{Tyr}$ gene sequence and ending at position 25. Tyr1 is perfectly matched to the $tRNA^{Tyr}$ gene sequence. Tyr2 contains a single nt mismatch [C→T] relative to position 16 of the $tRNA^{Tyr}$ gene.

B. Effect of temperature on hybridization of a M13 primer extension probe to human genomic DNA: A recombinant M13 DNA clone (~300 nt *Msp*I fragment from λHt7 - designated 52-4), containing a single $tRNA^{Lys}$ gene and flanking sequence (Doran et al., 1987), was used as template to prepare a probe using the M13 hybridization primer extension method. Hybridization of this probe to digested human genomic

DNA was performed at either 45° or 60°C. Membranes were washed at 50°C in 0.2x SSPE/0.1% SDS and were exposed to X-ray film for 3 days at -80°C using a Lightning Plus intensifying screen.

C. Effect of temperature on hybridization of random primer probes to human genomic DNA: A 610 bp *Eco*RI fragment from λHt4 (Roy et al., 1982) containing a human $tRNA^{Gln}$ gene and its flanking sequence was purified from low melting point agarose and used in the random primer probe preparation method. Approximately 5-30 ng DNA was used per labeling reaction. Hybridization to human genomic DNA (digested with *Bam*HI, *Bgl*II, *Eco*RI, or *Hin*dIII) was performed at 50° or 65°C in hybridization solution plus 7% SDS. Membranes were washed at either 50° or 65°C, respectively, and exposed to X-ray film for 3 days at -80°C using a Lightning Plus intensifying screen.

D. Effect of probe length: To investigate the effect of probe length on the hybridization of tRNA gene-containing probes to human genomic DNA, three different probes containing a single $tRNA^{Gln}$ gene and various amounts of flanking sequence were hybridized at 65°C to digested human genomic DNA. Hybridizations were performed in hybridization solution plus 10% dextran sulfate (w/v) for 16 hr. Dextran sulfate was used to increase the rate of hybridization of probes to rare sequences. To prevent excessive non-specific background sometimes associated with the use of dextran sulfate, the dextran sulfate was included in the prehybridization solution. Membranes were washed at 65°C in 0.2x SSPE/0.1% SDS and exposed to X-ray film for 3 days. The probes used for this experiment were (1) a 72 bp fragment containing a $tRNA^{Gln}$ gene without flanking sequence, (2) a 105 bp *Rsa*I/*Hin*dIII fragment from pHt4-10 (Roy et al., 1982) containing a $tRNA^{Gln}$ gene plus 33 bp of 5'-flanking sequence, and (3) a 610 bp *Eco*RI fragment from pHt4-10 containing a $tRNA^{Gln}$ gene plus 538 bp of flanking sequence.

E. Hybridization of random primer probes to mouse and human DNA: Approximately 10-20 μg of human genomic DNA or mouse genomic DNA (Balb/c, C3H) were digested with *Bam*HI, *Eco*RI, or *Hin*dIII and separated on a 0.8% agarose gel. The DNA was transferred to nylon membrane by the method of Rigaud et al., (1987). Probes were prepared by the random primer method. The DNA probes derived from various tRNA gene-containing human-λ recombinants are:

(1) λHtM2: a ~500 bp *Sty*I fragment containing a single $tRNA^{Tyr}$ gene plus flanking sequence isolated from pJM2 (MacPherson, 1988).

(2) λHtM4: a ~800 bp *ApaI*/*Hin*dIII fragment containing a single $tRNA^{Tyr}$ gene plus flanking sequence isolated from pJM4 (MacPherson, 1988). λHtM4 is known to contain two $tRNA^{Tyr}$ genes and a single $tRNA^{Ala}$ gene (D. Spadafora - unpublished).

(3) λHtM6: a ~600 bp *Eco*RI/*Sph*I fragment containing a single $tRNA^{Tyr}$ gene plus flanking sequence isolated from pJM6IT. λHtM6 is known to contain four $tRNA^{Tyr}$ genes (MacPherson and Roy, 1986).

(4) λHt4: a 610 bp *Eco*RI fragment encompassing a single $tRNA^{Gln}$ gene plus flanking sequence found within a tRNA gene heterocluster composed of single genes for $tRNA^{Lys}$, $tRNA^{Leu}$ and $tRNA^{Gln}$ (Roy et al., 1982).

(5) λHt7: a 545 bp fragment containing a single $tRNA^{Phe}$ gene plus flanking sequence found within a cluster of two $tRNA^{Phe}$ genes and two $tRNA^{Lys}$ genes isolated by PCR using the primers Phe1-(TTGCAGGGAACTTTAAAACAAC) and Phe2-(CACAGGGTGCTCAAGAAAGTG) and the template DNA pHt7-1 (Doran et al., 1987).

(6) λHt8: a ~5.2 kb *Bam*HI fragment containing a single $tRNA^{Gly}$ gene found on λHt8 (Doran et al., 1988) was cloned into pUC118. Four different sub-fragments containing the tRNA gene plus flanking sequence or only flanking sequence ranging in size from 235 bp to 515 bp derived by digestion of this plasmid with *Sma*I/*Sty*I, *Sst*II/*Sty*I, or *Sst*II/*Sma*I were used as probes.

(7) λHt9: a 275 bp fragment containing a single $tRNA^{Gly}$ gene plus flanking sequence was isolated by PCR using the primers Gly1-(CTCAACCTCTTACAA TCTCAGATAGG) and Gly2-(TGAAGATGTTTCTTAGAAATGCTACC) which were situated in the flanking sequences of this gene as found on λHt9 (Doran et al., 1988).

2.4.2 Mouse/human hybrid chromosomal mapping panels

*Bam*HI digested DNA from mouse/human somatic cell hybrids was fractionated by electrophoresis on 0.8% agarose gels and subsequently transferred to nylon membrane (Zeta probe/Dupont). These mapping panels were generously provided by Dr. Thomas Shows. The probes used to determine the chromosomal locations for some $tRNA^{Tyr}$ genes were as described previously. Fragments were purified from low melting point agarose (Boehringer Mannheim) and used directly in the random primer labelling method described by Feinberg and Vogelstein (1983, 1984). Approximately 10-30 ng of DNA was used per labeling reaction. All probes were purified by passage through a Sephadex G-50 column and were denatured at 94°C prior to use. Hybridization of probes to the chromosomal mapping panel was performed at 68°C for 16-24 hours in 10x Denhardt's solution, 6x SSPE, 3%SDS and ~100-500 μg sonicated *E. coli* DNA per ml hybridization solution. Membranes were washed twice at 65-68°C in 0.2x SSPE/0.1% SDS. Autofluorography was performed for 2-5 days using Kodak X-omat type AR film and a Dupont Lightning Plus intensifying screen(s) at -80°C. Chromosomal mappings of human tRNA gene-containing fragments were

performed as a double blind experiment. Chromosomal assignments were done by Dr. Thomas Shows.

2.5 Streptavidin-agarose selection of PCR amplification products

Approximately 10 µl of resuspended streptavidin conjugated agarose beads (Sigma: 1 ml packed gel of 4% cross-linked beaded agarose containing 1.2 mg of streptavidin per ml packed gel in 2 ml suspension) were washed with 100 µl Buffer A (10 mM Tris·HCl pH 8.0/1 mM EDTA/300 mM NaCl/100 µg/ml sonicated *E. coli* DNA), centrifuged at 16,000 rpm for 1 min and the supernatant was removed. Fifty microliters of each of the first stage PCR mixes was added to the washed beads. To this, 100 µl of Buffer A was added and the solution mixed by inversion for 10 min at room temperature. Each reaction was centrifuged as before and the supernatant was discarded. To the opaque agarose-streptavidin-biotin-DNA pellet was added 500 µl of Buffer B (10 mM Tris·HCl pH 8.0/1 mM EDTA/50 mM NaCl), the beads mixed by inversion several times and centrifuged as before. The supernatant was again discarded. These latter steps were repeated two to four times. After the final centrifugation step, the agarose beads were resuspended in 100 µl dH_2O. After mixing well, 10 µl of this selected DNA was then used for a second round of PCR as described for Alu-tRNA PCR (section 2.6.3) and Linker PCR (section 2.6.4).

2.6 Isolation of human tRNA gene sequences

2.6.1 tRNA PCR

PCR was carried out in a total volume of 100 µl using approximately 1 µg of total human genomic DNA. The tRNA gene-specific primers were used at a final concentration of ~2 µM in 70 mM Tris-HCl (pH 8.8 at 25°C); 2 mM $MgCl_2$; 0.1% Triton X-100; 250 µM each dNTPs, and 1 unit of Taq DNA polymerase (Boehringer Mannheim). Samples were overlayed with 50 µl of light mineral oil to prevent

evaporation. PCR was performed for 30 cycles of 94°C denaturation (30 sec), 58°C annealing (1 min), and 74°C extension (3 min). Primers were synthesized on an Applied Biosystems DNA synthesizer (model 381A or 391EP) and were used in reactions directly after deprotection and reprecipitation from 70% ethanol. Primer sequences used for PCR are as follows:

(1) $tRNA^{Tyr}$:	LTyr-CCTTCGATAGCTCAGCTGGTAGAG
	RTyr-TCCTTCGAGC(C/T)GGAAT(C/T)GAACAG
(2) $tRNA^{Phe}$:	LPhe-GCCGAAATAGCTCAGTTGGGAGAG
	RPhe-TGCCGAAACCCGGGAT(C/T)GTTCCTG
(3) $tRNA^{Lys}$:	LLys-GCCCGGATAGCTCAGTCGGTAGAGC
	RLys-CGCCCGAACAGGGACTTGAACCCTG
(4) $tRNA^{Trp}$:	LTrp-GACCTCGTGGCGCAACGGTAGC
	RTrp-TGACCCCGACGTGATTCGAACACG
(5) $tRNA^{Met}$:	LMet-GCCTGGTTAGCGCAGTAGGTAGCG
	RMet-TGGTGCCCCGTGTGAGGATCGAAC
(6) $tRNA^{Arg}$:	LArg-GTCTCTGTGGCGCAATCGGTT
	RArg-CGTCCCTGGGTGGGCTCGAAC

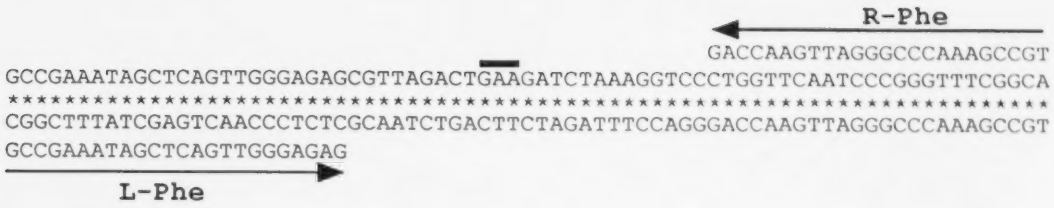
Mixed sites in the PCR primers represent positions of known polymorphisms.

PCR was done using each of the primers individually or with the other cognate tRNA gene-specific primer as indicated in Figure 8. After PCR, the mineral oil was removed by extracting with 50 µl chloroform. Amplification products were precipitated with ethanol and the DNA subsequently redissolved in 50 µl TE buffer (10 mM Tris-HCl, pH 8.0/ 1 mM EDTA). A 5 µl aliquot of each PCR product was used for electrophoresis in 2% agarose gels. DNA was transferred to nylon membrane as previously described and hybridized to one or both of the primers used in the original PCR to detect amplification products that contain the primer binding sites and also, to discriminate between those amplification products that contain a complete tRNA gene

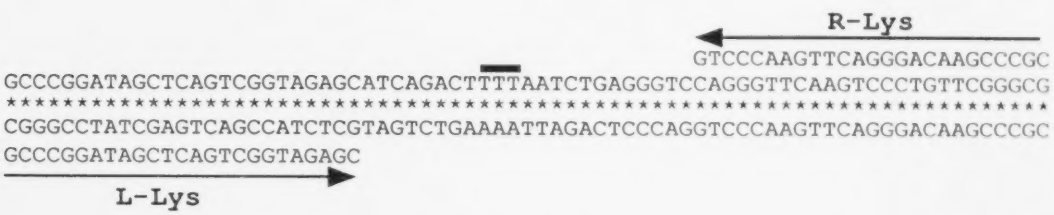
Figure 8 - tRNA PCR primer diagram.

This diagram shows the tRNA gene sequence for (a) $tRNA^{Phe}$, (b) $tRNA^{Lys}$, (c) $tRNA^{Tyr}$, (d) $tRNA^{Trp}$, (e) $tRNA^{His}$, (e) $tRNA_m^{Met}$, and (f) a $tRNA^{Arg}$ and the sequence of primers used for tRNA PCR for these tRNA gene families. Arrows indicate the orientation of each tRNA gene specific primer sequence (5' to 3') for each tRNA. Small solid lines indicate the anticodon region of each tRNA gene coding sequence. Each primer flanks the gene region that is expected to contain the intron sequence, if present (e.g. $tRNA^{Tyr}$). Known polymorphisms for tRNA or tRNA gene sequences are not shown. The possible tRNA gene sequences shown for $tRNA^{Trp}$, $tRNA^{His}$, and $tRNA^{Arg}$ are based on the sequences for the tRNAs isolated from mouse or bovine sources. A possible $tRNA_m^{Met}$ gene sequence is based on the human tRNA sequence. The 5' terminal G shown for $tRNA^{His}$ and the 3' terminal CCA shown for $tRNA^{Met}$ and $tRNA^{Arg}$ are added post-transcriptionally. Sequences are from Sprinzl et al., 1987.

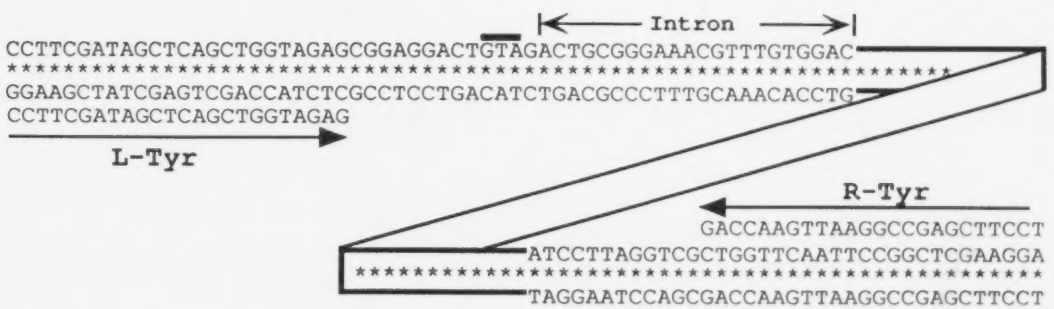
a) tRNA ^{Phe}



b) tRNA ^{Lys}



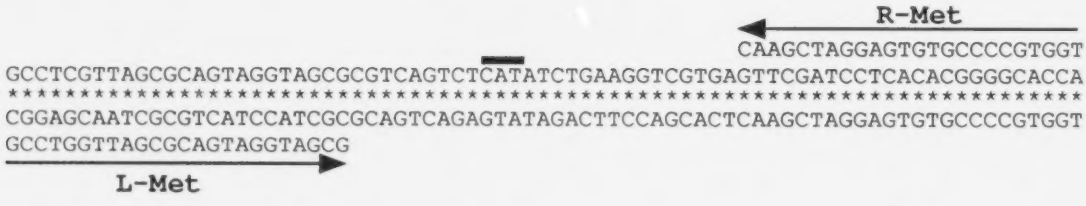
c) tRNA ^{Tyr}



d) tRNA^{Trp}

e) tRNA^{His}

f) tRNA^{Met}



g) tRNA^{Arg}

sequence and those which contain tRNA-like sequences or partial tRNA gene sequences. Human-pUC118 recombinants which contain DNA fragments derived from tRNA PCR for the $tRNA^{Phe}$ and $tRNA^{Tyr}$ gene families were given the root designations pHf (pHf1-39; pHf7-30; pHf7-31) and pHy (pHy6; pHy8; pHy9; pHy10; pHy15; pHy17; pHy20), respectively.

To determine if the various fragments isolated by tRNA PCR for the $tRNA^{Tyr}$ gene family were artifacts, PCR was performed using primers that were specific for the intron sequences found in pHy6, pHy10, pHy15 and pHy20, and a second primer that was common to the intergenic flanking region. The sequences of the primers used are:

Tyr3-TGTAGCTACTTCCTCAGCAGGA (pHy10/20 introns)

Tyr4-TGTAGATTGTACAGAAATTTGC (pHy6 intron)

Tyr5-ATGTCTCCTGCTGAGGAAGTAG (pHy10 intron)

Tyr6-ACTGTAGGGGTTTGAATGTGGC (pHy15 intron)

Tyr7-CGATGACTTATGGCACTTTCC (intergenic primer)

Tyr8-CGCCTTGTCTCCATAACATTG (intergenic primer)

Tyr9-CGGGCCCGAGTCACACAGGA (intergenic probe)

PCR was done under standard conditions, using either human genomic DNA or plasmid DNA as the template. Approximately 1 µg (human DNA) and 1 ng (plasmid DNA) were used. The combinations of primers and template DNA used in PCR are: (1) Tyr3+Tyr7/human DNA, (2) Tyr3+Tyr7/pHy20 DNA,(3) Tyr4+Tyr7/human DNA, (4) Tyr4+Tyr7/pHy6 DNA, (5) Tyr5+Tyr7/human DNA, (6) Tyr5+Tyr7/pHy15 DNA, (7)Tyr6+Tyr8/human DNA, and (8)Tyr6+Tyr8/pHy10 DNA. Twenty microliters from each reaction was fractionated on a 1% agarose gel. DNA was transferred to nylon membrane and hybridized to end labelled Tyr9 DNA.

2.6.2 Inverted PCR

Approximately 1 μg human genomic DNA was digested with *Cfo*I, *Hae*II, *Hin*PI, *Hpa*II or *Hin*dIII and then diluted serially 10-fold from 10^0 to 10^{-5} in 10 μl final volume. Five microliters of each dilution was used in a ligation reaction using 1 unit T4 DNA ligase for ~16 hr at room temperature (20 μl final reaction volume). Ten microliters of each ligation reaction was used in a PCR, as previously described for tRNA PCR, using primers that were specific for the conserved regions of $tRNA^{Phe}$, $tRNA^{Trp}$, or $tRNA^{His}$ genes. The sequence of the primers used for inverted PCR are:

(1)$tRNA^{Phe}$ IL-Phe: CTCCCAACTGAGCTATTTCGGC,

IR-Phe: AAGATCTAAAGGTCCCTGGTTC,

(2)$tRNA^{Trp}$ IL-Trp: GCGCTACCATTGCGCCACGAGGT,

IR-Trp: CTCCAGATCAGAAGGTTGCGTGTTC,

(3)$tRNA^{His}$ IL-His: GTACTAACCACTATACGATCACGGC,

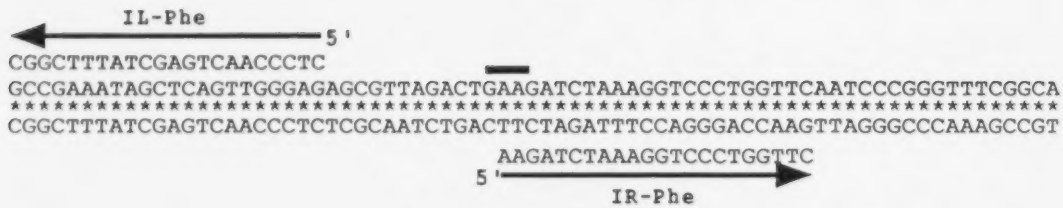
IR-His: TTGTGGCCGCAGCAACCTCGGT,

The 5' gene-specific primer for each tRNA gene family starts in the distal region of the dihydrouridine stem and ends at the 5' terminus of the tRNA gene. The 3' gene-specific primer starts around position 33 of the anticodon loop region and ends in the pseudouridine loop (see Figure 9). Primers were oriented in opposite directions and should only serve as amplification primers if the template DNA was circularized during the ligation step. Amplification products of tRNA genes using this arrangement of primers would have the 3' terminal conserved gene sequence incorporated as directed by the template DNA and would thus serve as means of identifying true tRNA gene amplification. These inverted primers would also serve as PCR primers if the tRNA genes occurred as tandem repeats.

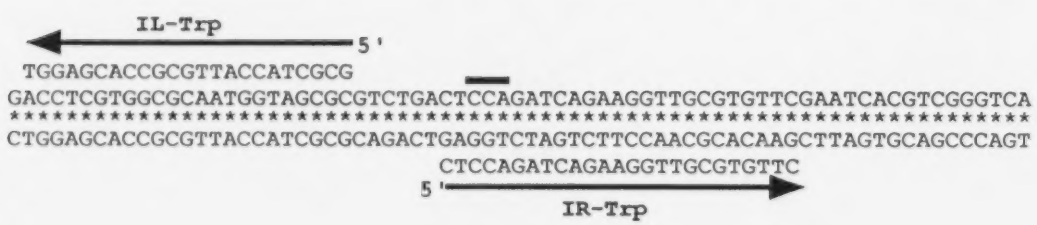
Figure 9 - Inverted PCR primer diagram.

This diagram shows the tRNA gene coding sequences for (a) $tRNA^{Phe}$, (b) $tRNA^{Trp}$, and (c) $tRNA^{His}$ and the sequence of primers used for inverse PCR for these tRNA gene families. Arrows indicate the orientation of each tRNA gene-specific primer sequence (5' to 3') for each tRNA. Small solid lines indicate the anticodon region of each tRNA gene coding sequence. The regions between the 5' ends of each primer should not be included in the amplification products. Additionally, the 3' gene-specific primer does not include the last ~16 nt of coding sequence, which should serve as a control for amplification from complete tRNA genes. A probe specific for the terminal 3' region was used to hybridize to amplification products from PCR using these primers. Any fragments containing this region of the tRNA gene must be derived from the template encoded gene sequence. Sequences are from Sprinzl et al., 1987.

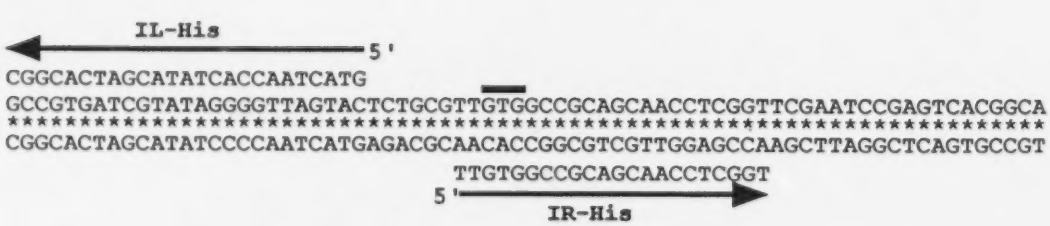
a) tRNA^{Phe}



b) tRNA^{Trp}



c) tRNA^{His}



After PCR, the mineral oil was removed and the DNA ethanol precipitated. The DNA was redissolved to 100 µl in TE buffer. Ten microliters was used for electrophoresis on 1.5% agarose gels and the DNA transferred to nylon membranes. DNA from each set of PCRs was hybridized with an oligonucleotide which was specific for a region of the tRNA sequence that was not contained in either of the primer sequences used for the PCR (e.g. RPhe, RTrp, RHis). Amplification products from inverted PCR for the $tRNA^{Phe}$ gene family (*Cfo*I, *Hin*PI, and *Hae*II samples) were cloned into the *Hin*cII site of pUC118 (donated by J. Vieira). Recombinants were subsequently screened by colony hybridization with 5'-end labelled RPhe (see section 2.6.1-tRNA PCR for oligonucleotide description). Positively hybridizing recombinants were used for sequence analysis. The plasmid designation for recombinants obtained from these experiments was pCfo, pHae and pHnp.

In order to reconstitute the correct orientation of DNA sequences as contained in the human genome, PCR was performed on human DNA, using primers which were oriented in non-overlapping directions based on the sequence for pHnp4 and pHnp12. These recombinants contain fragments from inverted PCR for the $tRNA^{Phe}$ gene family which exhibit the restriction recognition sequence [G↓CGC] of the enzyme (*Hin*PI) used in the original digestion step. Primers were positioned flanking the restriction recognition site. The primers used for each PCR are:

Hnp4-L:CAGTGATAGTTTACTCAGGAG

Hnp4-R:TCGCTCCTGGGGGTGTGTTAG

Hnp12-L:ATCCCACCTACTGGAAGAGAC

Hnp12-R:GCTGTGTTGAAGGAGTGACAA

Amplification products were fractionated on a 1.5% agarose gel and transferred to nylon membrane. This was hybridized with 5'-end labelled IR-Phe (which is specific for the 3' region of the $tRNA^{Phe}$ gene sequence).

2.6.3 Alu-tRNA PCR

A two stage PCR using Alu repeat sequence-specific primers and tRNA gene-specific primers for $tRNA^{Phe}$ and $tRNA^{Lys}$ genes was performed following the PCR protocol previously described for tRNA PCR except that the extension time at 72°C-74°C was increased to 5 min.

The $tRNA^{Phe}$ gene-specific primers are as follows:

b-LPhe: biotin - C_6 linker - (GCCGAAATAGCTCAGTTGGGAGAGCTT AGACTGAAGATC)

LPhe: (GCCGAAATAGCTCAGTTGGGAGAG)

b-RPhe: biotin - C_6 linker - (TGCCGAAACCCGGGATCGAACCAG)

RPhe: (TGCCGAAACCCGGGATCGAACCAG)

The $tRNA^{Lys}$ primers are as follows:

b-LLys: biotin - C_6 linker - (GCCCGGATAGCTCAGTCGGTAGAGCAT CAGACTTTTAATC)

LLys: (GCCCGGATAGCTCAGTCGGTAGAGC)

b-RLys: biotin - C_6 linker -(CGCCCGAACAGGGACTTGAACCCTG)

RLys: (CGCCCGAACAGGGACTTGAACCCTG)

The Alu specific primers are:

Alu1: (CTCCCAAAGTGCTGGGATTACAGG)

Alu2: (CCTGTAATCCCAGCACTTTGGGAG)

Alu3: ([C/T][G/A]CCA[C/T]TGCACTCCAGCCTGGG)

Alu4: ([C/T][G/A]CCCAGGCTGGAGTGCAG/ATGG)

The non-biotinylated oligonucleotides were synthesized on an Applied Biosystems DNA Synthesizer model 381A or 391EP and were used directly after synthesis without chromatographic purification. Biotinylated primers were purified by passage through a Sephadex G-50 column and elution in low salt conditions.

In the first stage PCR, Alu-specific primers and biotinylated tRNA gene-specific primers were used with human genomic DNA as the template. Primer

combinations used were as follows: (1)b-LPhe + Alu1, (2)b-LPhe + Alu2, (3)b-LPhe + Alu3, (4)b-LPhe + Alu4, (5)b-LLys + Alu1, (6)b-LLys + Alu2, (7)b-LLys + Alu3, (8)b-LLys + Alu4. Additionally, control PCR reactions with only one Alu primer (Alu1/Alu2/ Alu3/Alu4) were performed to check for non-specific amplification of tRNA sequences contained between two Alu sequences. Samples were overlaid with 50 µl light mineral oil to prevent evaporation and PCR was performed for 30 cycles at 94°C (30 sec): 56°C (1 min): 72°C (5 min).

After completion of the first PCR step, the samples were extracted with 50 µl chloroform and each sample ethanol precipitated. The amplification products were dissolved in 100 µl dH_2O, of which 50 µl was saved for further characterization, and the remaining 50 µl used for the streptavidin selection step and the subsequent second round of PCR using non-biotinylated primers. The primers used for this PCR were similar to the first set except that non-biotinylated primers were used (e.g. LPhe [b-LPhe], LPhe [b-LPhe], LLys [b-LLys], LLys [b-LLys]) to allow possible cloning of the amplification products. The parameters for this round of PCR were the same as those used for the first round. After PCR, each sample was extracted with 50 µl chloroform and the samples were precipitated with ethanol. Each sample was redissolved to 100 µl with TE buffer (10 mM Tris·HCl pH 8.0/1 mM EDTA). A 10 µl aliquot of each set of PCR products (unselected vs selected) was fractionated by electrophoresis on a 2% agarose gel. DNA was transferred to nylon membranes for hybridization with the 3' gene-specific oligonucleotides.

2.6.4 Linker PCR

Human genomic DNA was digested with restriction enzymes that generate blunt ends (e.g. *Alu*I, *Eco*RV, *Hae*III, *Hpa*I, *Rsa*I, *Sma*I, or *Ssp*I). The oligonucleotides Plinker1-(GAAGCTTGGATCCGTCGACGAATTC) and Plinker2-(GAATTCGTCGACGGATCCAAGCTTC) were heated to 94°C for 3 min then slowly

cooled to room temperature to allow annealing of these complementary oligonucleotides. Ligation of linkers was performed using 1 unit of T4 DNA ligase for 16 hr at room temperature at an estimated molar ratio of 5:1 of the annealed linkers to the estimated number of DNA ends based on the average length of fragments generated for a given restriction digest of human genomic DNA. Amplification from human genomic DNA was done using a tRNA gene-specific primer and a second primer specific to a primer binding site that was provided by the ligation of linkers. In order to provide additional specificity of amplification, a two stage PCR involving biotinylated tRNA gene-specific primers was performed. The biotinylated amplification products were selected using streptavidin-agarose and a second PCR done using either a non-biotinylated tRNA gene-specific primer with a sequence similar to the biotinylated primer or with a nested primer (i.e. the same primers used for inverted PCR for a given tRNA gene family) that was specific to a region internal to the first tRNA gene-specific primer.

The first stage PCR was performed using the primer combinations: (1) b-LPhe + Plinker1, and (2)b-RPhe + Plinker1. Each reaction was extracted with 50 µl chloroform and the amplification products ethanol precipitated. DNA was then redissolved to 100 µl in TE buffer of which 50 µl was used for streptavidin selection. After streptavidin selection, a second round of PCR was performed using the appropriate amplification reactions from the first round PCR and the primer combinations: (1) LPhe + Plinker1 and (2) RPhe + Plinker1. For those second stage PCR reactions involving nested primers, the following primer combinations were used: (1) IL-Phe + Plinker1, and (2) IR-Phe + Plinker1. The nested primer IL-Phe was used on those PCR amplification products generated using RPhe whereas the nested primer IR-Phe was used on amplification products generated from PCR using LPhe.

2.7 Cloning and screening of PCR amplification products

DNA segments generated by PCR were cloned directly into the *Hin*cII site of pUC118. Recombinant plasmids were transformed into *Escherichia coli* strains JM83 or MV1193. Plasmid containing cells were grown and selected on 2x YT + ampicillin (~100 mg/L) agar plates (Vieira and Messing, 1987). Initial screening for recombinant plasmids was performed by growing the cells in the presence of IPTG and X-gal. The presence of white colonies was indicative of an insert that disrupted the lacZ' gene. Screening of large numbers of transformants for tRNA gene-containing inserts was done using a modified colony lift procedure (Grunstein and Hogness, 1975). Duplicate plates of recombinant DNA-containing cells were grown on 2x YT+ amp plates and on nylon membranes (Hybond 'N'/Amersham) placed on 2x YT + amp plates. After growing at 37°C overnight, cells were lysed directly on the nylon membranes by placing the membranes in 0.5 M NaOH for 5-10 minutes, two washes in 1 M Tris-HCl buffer (pH 8.0) for 2 minutes, once in 0.5 M Tris-HCl, pH 8.0; 1.5 M NaCl for 3 minutes and finally the membranes were washed in 2x SSPE for 15 minutes. Cell debris was carefully removed and the membranes were baked under vacuum at 80°C for 2 hr. Screening for recombinants which contain inserts was done by hybridizing either one or both of the PCR primers to the membranes. Alternatively, when suitable primers existed that were contained within the amplifiable DNA segment but found outside the regions covered by the PCR primers, these were used as hybridization probes. Positively hybridizing transformants were picked and grown overnight at 37°C in 3 ml 2x YT + ampicillin broth. Plasmid DNA was isolated using the alkaline lysis method of Birnboim (1983). DNA was fractionated electrophoretically on 0.75% agarose gels, transferred to nylon membrane and again hybridized to tRNA gene-specific probes.

2.8 DNA sequencing

The complete sequence of tRNA containing clones was determined in both directions by double stranded sequencing (Chen and Seeburg, 1985) using the dideoxy chain termination method as described by Sanger et al., (1977). Radioactive sequencing products were separated on 6% polyacrylamide gels (38:2, acrylamide:N, N'-methylene bisacrylamide; 8.3 M urea) using TEB running buffer (60 mM Tris, 60 mM boric acid, and 1.2 mM EDTA). Autoradiography of labelled DNA bands was performed for 8-16 hrs at -20°C using Kodak X-omat AR film. Alternatively, double stranded plasmid DNA was sequenced using fluorescent dye-terminators according to the manufacturer's suggestions (Applied Biosystems Ltd.). Fluorescent sequencing products were then separated and analysed on the Applied Biosystems model 373A DNA sequencer. Sequencing results were further analyzed using the DNAstar DNA analysis software (DNAstar Inc.)

3. RESULTS

3.1 Standardization of hybridization conditions

***3.1.1. Oligonucleotide hybridization*:** No signal was detected when the Tyr1 DNA probe was hybridized to human genomic DNA at 45° or 48°C, . However, when the hybridization was repeated at 42°C, faint bands were visible after an exposure time of 5 days to X-ray film (data not shown). Direct examination of the autoradiograph showed that at least 12 bands are visible in the *Eco*RI digests of human DNA, while a lesser number is visualized in the other lanes. There was no detectable hybridization of an oligonucleotide (Tyr2), containing a single nt mismatch (C→T) relative to position 16 of the $tRNA^{Tyr}$ gene sequence, under any conditions examined. The single nt difference between Tyr1 and Tyr2 was positioned at the center of the oligonucleotides to maximize the disruptive effect on hybridization to mismatched sequences within the genome. These probes were designed to allow competitive hybridization between the two oligonucleotides (Tyr1 and Tyr2) to allow specific detection of $tRNA^{Tyr}$ genes which contained this polymorphism. Alternatively, diminished specific hybridization of Tyr2 in the presence of unlabelled Tyr1 might be interpreted to mean that this polymorphism seen in the tyrosine tRNAs (at position 16) is due to post-transcriptional modification. However, the weak hybridization results seen for both probes prevents conclusions to be drawn. Moreover, while the detection of $tRNA^{Tyr}$ gene-containing fragments in human genomic DNA was possible under these conditions, the low signal seen prevents the use of oligonucleotide probes for the detection of specific loci in mouse-human somatic cell hybrid DNA preparations, where the number of human target sequences would be considerably reduced.

***3.1.2. M13 Primer Extension*:** The M13 primer extension probe preparation method was examined to determine whether the specific activity of radiolabelled tRNA gene-containing DNA fragments could be increased to lessen the exposure time

required for the detection step. Figure 10 shows the results of hybridization of a ~400 bp $tRNA^{Lys}$ gene-containing probe (*Msp*I fragment from λHt7; Doran et al., 1987) to human genomic DNA at 45° and 60°C. At 45°C, there is significant background hybridization that tends to obscure the specific signal seen due to hybridization of this probe to the 8.8 and 9.4 kb bands seen in *Eco*RI and *Hin*dIII digests, respectively. The background smear seen under these conditions may be the results of hybridization of repeated sequences contained within the probe insert sequence or, alternatively, it may be due to sequences in M13 DNA that are complementary to repeated sequences found within the human genome (Vassart et al., 1987). When the temperature of hybridization was increased to 60°C, the non-specific hybridization signal decreased drastically. By increasing the stringency of hybridization for this probe it was possible to detect specific sequences in human genomic DNA preparations, however the specific activity of the probe appeared not to be sufficient to allow detection of a strong positive signal within a reasonable time. Intermediate temperatures may have yielded better results, but these were not examined.

3.1.3. Random Primer Extension Method: The random primer probe preparation method of Feinberg and Vogelstein (1983, 1984) was subsequently examined. Samples of DNA labelled by random primer extension were passed through a Sephadex G-50 column, and 1 µl aliquots were removed for scintillation counting. The specific activity of probes labelled using this method was estimated to be 1-5 x 10^8 cpm/µg, assuming 30 ng of DNA was used per reaction.

Figure 11 shows the hybridization of a 610 bp $tRNA^{Gln}$ gene-containing probe (from λHt4: *Eco*RI fragment; Roy et al., 1982) to human genomic DNA at 50° and 65°C. Again, as was seen with the M13 primer extension method, hybridization at the less stringent temperature of 50°C resulted in a significant amount of background hybridization, however the intensity of the specific hybridization was noticeably

Figure 10 - Hybridization of a human-M13 recombinant (52-4) to human genomic DNA.

Hybridization was performed at 45°C or 60°C to DNA digested with *Eco*RI (lane A) or *Hin*dIII (lane B). The M13 recombinant 52-4 contains a ~400 bp *Msp*I fragment, encompassing a $tRNA^{Lys}$ gene found on pHt7-1 (Doran et al., 1987). Probe DNA was prepared by the M13 primer extension method. Membranes were washed at the same temperature at which hybridization was performed and were exposed to X-ray film for 3 days using a Lightning Plus intensifying screen.

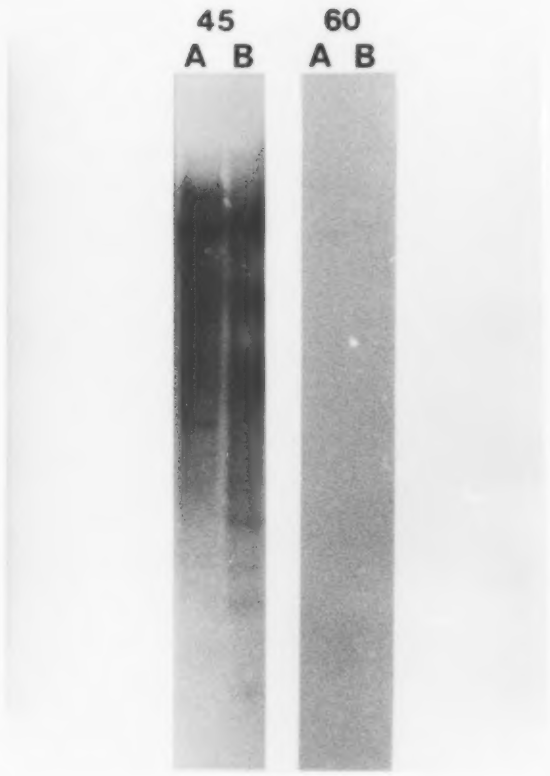
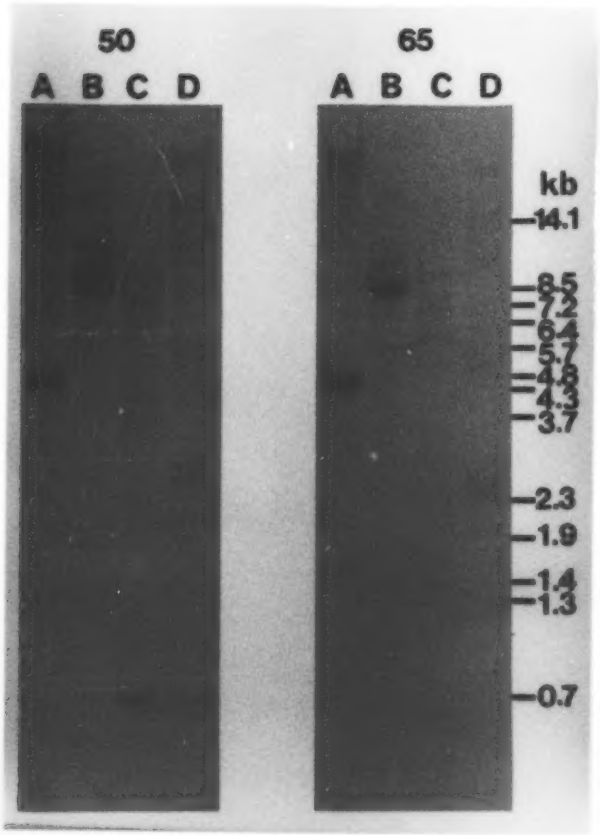

45
A B
60
A B

Figure 11 - Hybridization of a human $tRNA^{Gln}$ gene probe to human genomic DNA.

Hybridization was performed at 50° or 65°C to DNA digested with *Bam*HI (lane A), *Bgl*II (lane B), *Eco*RI (lane C), or *Hin*dIII (lane D). The probe used for these hybridizations was a 610 bp *Eco*RI fragment, containing a single $tRNA^{Gln}$ gene and its flanking sequence, isolated from λHt4. Probe DNA was prepared by random primer extension. Membranes were washed at the same temperature at which hybridization was performed and were exposed to X-ray film for 3 days using a Lightning Plus intensifying screen.

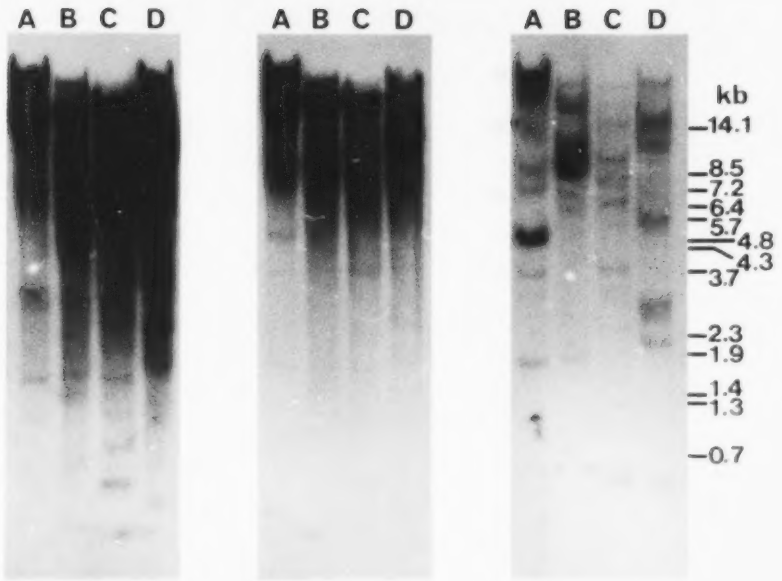

50
A B C D
65
A B C D
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

stronger. When the temperature of hybridization was increased to 65°C the background hybridization was significantly reduced while a relatively strong specific signal was still detectable in two of the four lanes.

The effect of probe length/composition on hybridization is shown in Figure 12. From this it can be seen that the hybridization to specific sequences within the human genome can be increased by increasing the probe length (thus increasing the amount of unique flanking sequence contained within the probe). A 72 bp probe containing only the gene coding sequence for $tRNA^{Gln}$ exhibited a strong degree of hybridization to alternate sequences within the human genome as compared to the hybridization of a 610 bp probe containing the $tRNA^{Gln}$ gene plus the flanking sequence surrounding this gene. The size of the fragments showing the strongest hybridization intensity in the *Bam*HI and *Bgl*II digests of human DNA when probed with the 610 bp probe agree with that expected based on the restriction map for λHt4. Hybridization to the expected fragment (610 bp) in the *Eco*RI digests is only weakly detectable, as is also observed for the hybridization to human DNA digested with *Hin*dIII. Differences in results for the hybridization of this probe to genomic DNA shown in Figure 11 and Figure 12 are probably due to the use of dextran sulfate in the latter hybridization. A 105 bp probe containing the $tRNA^{Gln}$ gene and a smaller subsection of the 5'-flanking sequence also demonstrated a considerable amount of non-specific hybridization. Importantly, the most intense bands visualized when hybridized with these three similar probes was not the same. Presumably, the increased specificity of hybridization for the longer probe is due to the effect of the unique flanking sequence. From these results it was decided that the longer probes were most suitable for the purposes of hybridization to chromosomal mapping panels.

Figure 12 - Effect of flanking sequence on the hybridization of human $tRNA^{Gln}$ gene specific probes to human genomic DNA.

DNA was digested *Bam*HI (lane A), *Bgl*II (lane B), *Eco*RI (lane C), or *Hin*dIII (lane D). The left panel shows the results of hybridization for a 72 bp DNA fragment containing only the human $tRNA^{Gln}$ gene sequence. The middle panel shows the results of hybridization for a 105 bp *Rsa*I-*Hin*dIII fragment containing a human $tRNA^{Gln}$ gene plus 33 bp of 5' flanking sequence isolated from λHt4. The right panel shows the results of hybridization for a 610 bp *Eco*RI fragment containing a single human $tRNA^{Gln}$ gene plus 538 bp of flanking sequence (5' + 3') isolated from λHt4. Probe DNA was labelled by random primer extension. Hybridizations were performed at 65°C in hybridization solution plus 10% dextran sulfate (w/v). Membranes were washed at ~65°C and were exposed to X-ray film for 3 days in the presence of a Lightning Plus intensifying screen.


A B C D
A B C D
A B C D
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

3.1.4. Hybridization to Mouse and Human Genomic DNA: The greatest proportion of DNA in the chromosomal mapping panels is of mouse origin. It was therefore necessary to determine whether the probes used for the detection of human tRNA genes would detect sequences within the mouse genome, and if similar sequences were present, whether restriction fragment length polymorphisms would allow the distinction of hybridization to human or mouse sequences.

Genomic DNA isolated from human and mouse sources was digested with various restriction enzymes and was size fractionated on 0.8% agarose gels. After transferring the DNA to nylon membranes, the blots were hybridized under very stringent conditions (e.g. 68°C hybridization/65-68°C wash in 0.2x SSPE) with probes specific for human tRNA gene loci. The results of these hybridizations are shown in Figures 13 to 19. In all cases it was possible to discern differences in hybridization between the human and mouse DNA samples. In those examples where there were significant signals observed in the mouse DNA lanes, the signals were associated with different lengths of DNA, thus making determination of the human-specific hybridization event possible. This was most noticeable for the λHt4 specific probe that contains a single glutamine tRNA gene (linked to genes for $tRNA^{Lys}$ and $tRNA^{Leu}$). The strong specific hybridization to mouse genomic DNA observed for this probe may mean that there has been very little evolutionary divergence at this locus between human and mouse DNAs. Alternatively, multiple copies of a homologous DNA segment may be present in the mouse genome. This fragment of human DNA had previously exhibited unusual hybridization as observed in a study of the effect of probe length (see Figure 12). This probe also demonstrated extremely strong hybridization to lambda DNA (at 50°C plus 50% formamide) when care was taken to ensure that the probe DNA was not contaminated with vector DNA (data not shown). While it may be possible that the probe DNA was contaminated with vector DNA that was homologous to some lambda sequences, it is also possible that the tRNA gene or

Figure 13 - Hybridization of a λHtM2 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a ~500 bp *Sty*I fragment, containing a single human $tRNA^{Tyr}$ gene plus flanking sequence isolated from λHtM2, to human genomic DNA (lanes A1, B1, C1), Balb/C mouse genomic DNA (lanes A2, B2, C2) or C3H mouse genomic DNA (lanes A3, B3, C3) was done at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~3 days in the presence of a Lightning Plus intensifying screen at -80°C.

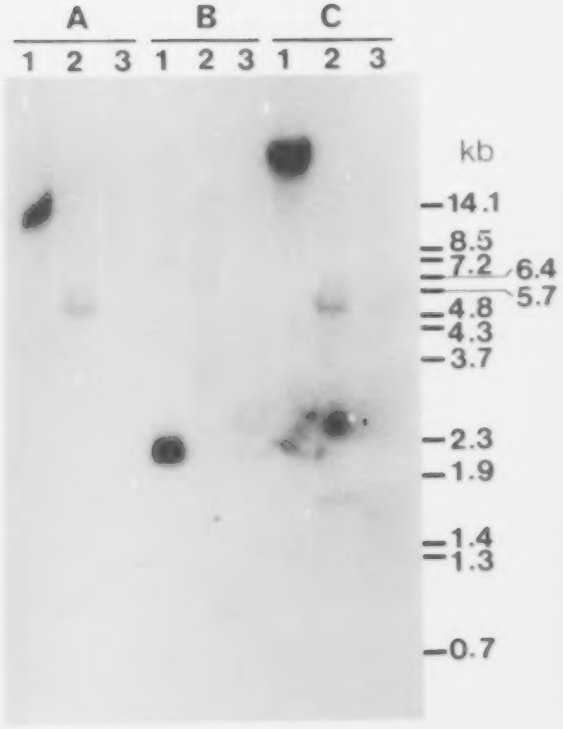

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 14 - Hybridization of a λHtM4 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a ~800 bp DNA fragment, containing a single human $tRNA^{Tyr}$ gene plus flanking sequence isolated from λHtM4, to human genomic DNA (lanes A1, B1, C1), Balb/C mouse genomic DNA (lanes A2, B2, C2) or C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~3 days in the presence of a Lightning Plus intensifying screen at -80°C.

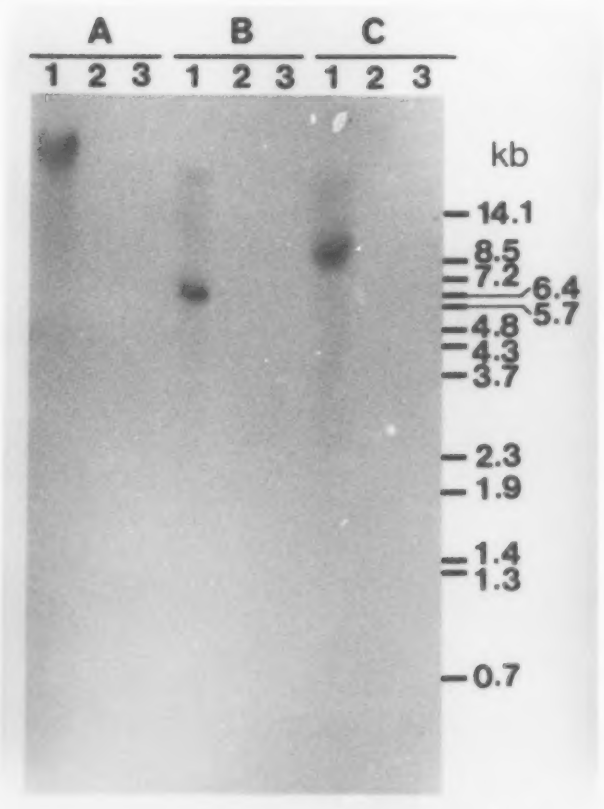

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 15 - Hybridization of a λHtM6 ($tRNA^{Tyr}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a ~600 bp *Eco*RI/*Sph*I DNA fragment, containing a single human $tRNA^{Tyr}$ gene plus flanking sequence isolated from λHtM6, to human genomic DNA (lanes A1, B1, C1), Balb/C mouse genomic DNA (lanes A2, B2, C2) or C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~3 days in the presence of a Lightning Plus intensifying screen at -80°C.

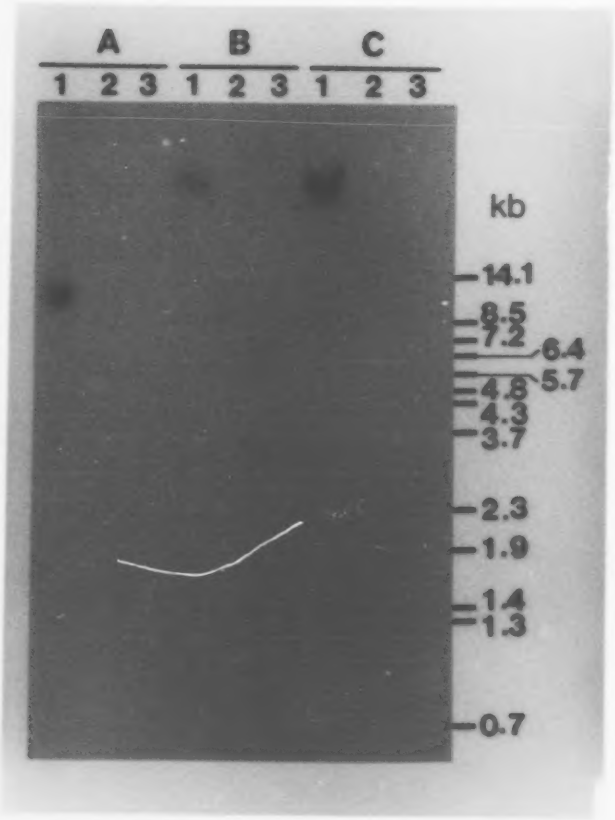

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 16 - Hybridization of a λHt4 ($tRNA^{Gln}$, $tRNA^{Leu}$, $tRNA^{Lys}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a 610 bp *Eco*RI DNA fragment, containing a single human $tRNA^{Gln}$ gene plus flanking sequence isolated from λHt4, to human genomic DNA (lanes A1, B1, C1), Balb/C mouse genomic DNA (lanes A2, B2, C2) or C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~4 days in the presence of a Lightning Plus intensifying screen at -80°C.

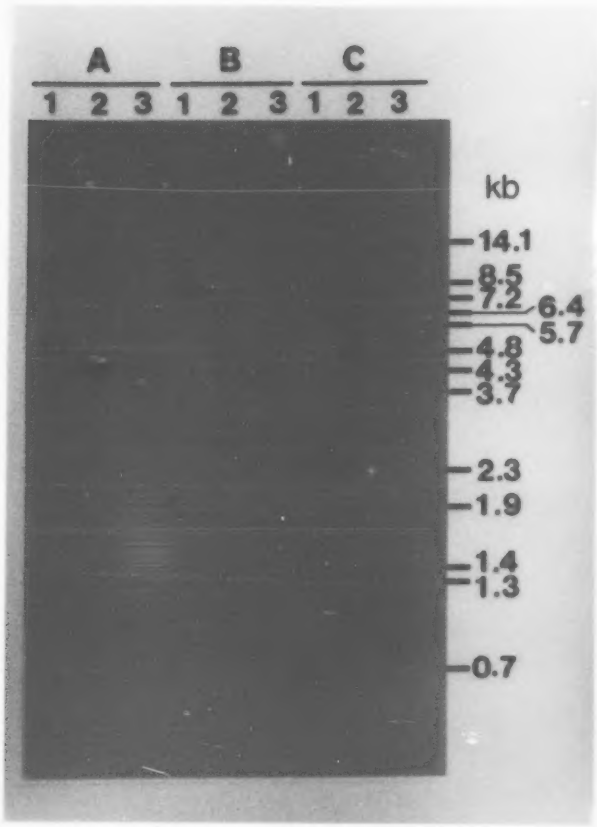

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 17 - Hybridization of a λHt7 ($tRNA^{Phe}$, $tRNA^{Lys}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a ~300 bp DNA fragment, containing a single human $tRNA^{Phe}$ gene plus flanking sequence isolated from λHt7, to human genomic DNA (lanes A1, B1, C1), Balb/C mouse genomic DNA (lanes A2, B2, C2), or C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~4 days in the presence of a Lightning Plus intensifying screen at -80°C.

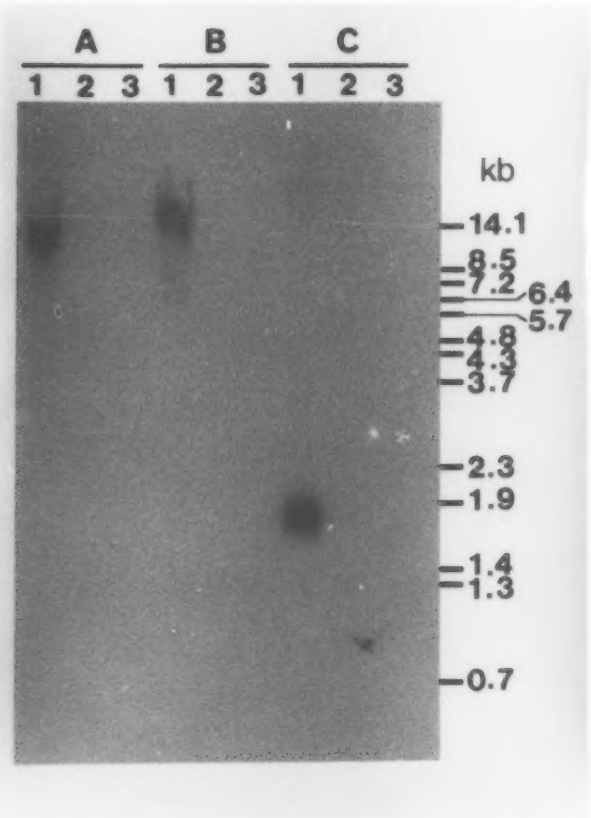

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 18 - Hybridization of a λHt8 ($tRNA^{Gly}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a 545 bp DNA fragment, containing a single human $tRNA^{Gly}$ gene plus flanking sequence isolated from λHt8, to human genomic DNA (lanes A1, B1, C1), to Balb/C mouse genomic DNA (lanes A2, B2, C2) or to C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~2 days in the presence of a Lightning Plus intensifying screen at -80°C.

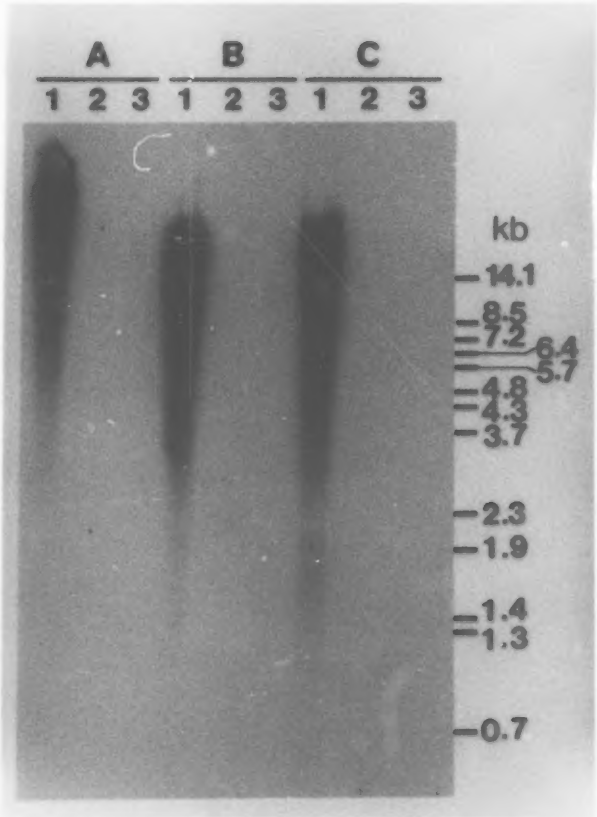
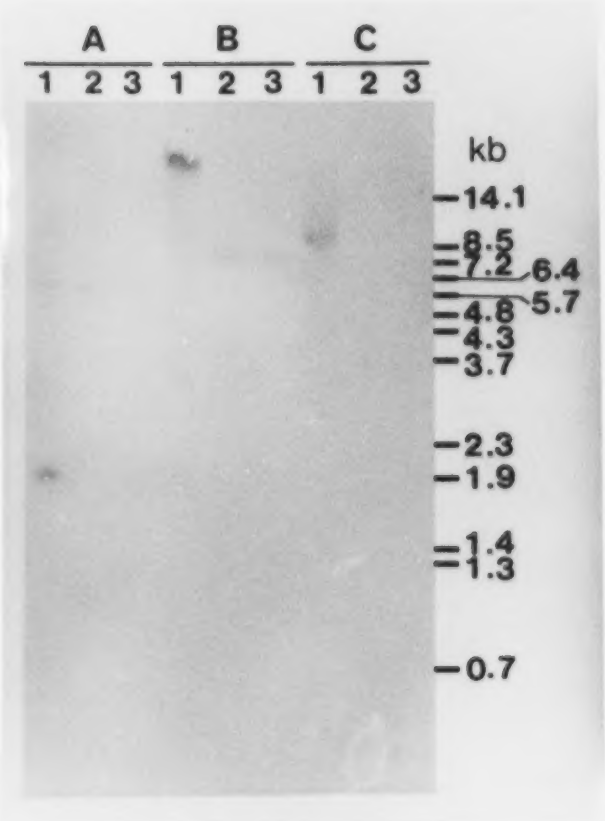

A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Figure 19 - Hybridization of a λHt9 ($tRNA^{Gly}$) specific probe to human and mouse genomic DNA.

All DNA samples were digested with *Bam*HI (A), *Eco*RI (B), or *Hin*dIII (C). Hybridization of a ~300 bp DNA fragment, containing a single human $tRNA^{Gly}$ gene plus flanking sequence isolated from λHt9, to human genomic DNA (lanes A1, B1, C1), to Balb/C mouse genomic DNA (lanes A2, B2, C2) or to C3H mouse genomic DNA (lanes A3, B3, C3) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. The membrane was washed twice at ~65-68°C in 0.2x SSPE, 0.1% SDS washing solution and was then exposed to X-ray film for ~4 days in the presence of a Lightning Plus intensifying screen at -80°C.


A
B
C
1 2 3 1 2 3 1 2 3
kb
14.1
8.5
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

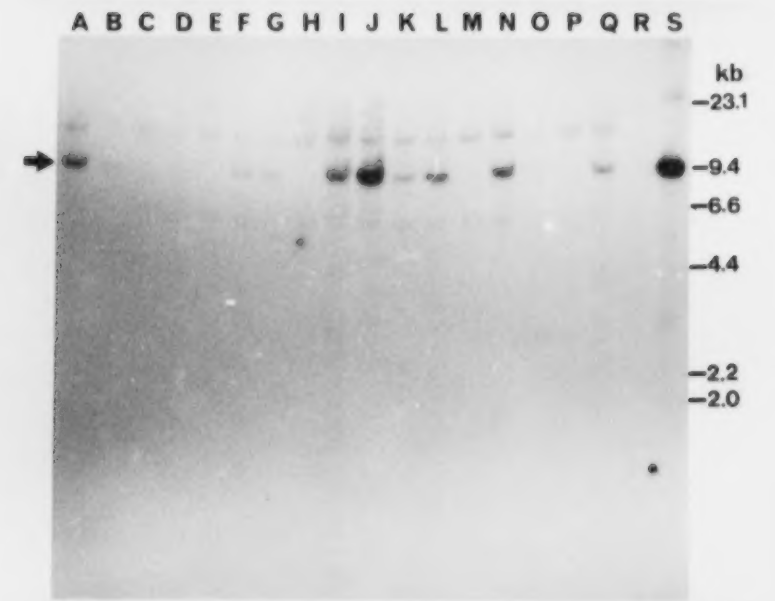
flanking sequences are homologous to some sequences contained in the lambda DNA. Hybridization of the λHt8 probe ($tRNA^{Gly}$) resulted in strong signals throughout the lane for the human DNA and no signal present for the mouse DNA. This is probably due to the presence of a human repeat element in the probe sequence. In order to overcome this problem, this fragment was digested with *Sma*I/*Sty*I, *Sst*II/*Sty*I or *Sst*II/*Sma*I to generate four fragments that span a ~600 bp region containing this glycine tRNA gene. These fragments were isolated from low melting point agarose and used as probe DNA for the random primer probe preparation method. In all cases, there was the same hybridization pattern observed as that shown in Figure 18 (data not shown). Thus, for this particular tRNA gene, detection of a specific band in the human DNA lane was not observed. Competitive hybridization using unlabelled human DNA may eliminate much of the background hybridization to human repeat sequences observed for these probes. Better results could also be achieved by isolating DNA from some other genetically linked region (i.e. some other fragment from λHt8) for use as a probe.

3.2 Human tRNA gene chromosomal localization studies

Those human tRNA gene-specific probes that exhibited strong specific hybridization to the human plus mouse DNA panels were used as probes for hybridization to a mouse/human somatic cell hybrid chromosomal mapping panel. This chromosomal mapping panel consisted of DNA from 34 different hybrids digested with *Bam*HI (two membranes) that contain various assortments of human chromosomes amongst the background of the mouse genome. The results of hybridization of these probes are shown in Figures 20 to 23. A ~9.4 kb fragment was observed to hybridize to the λHtM6 specific probe ($tRNA^{Tyr}$). While there was a very strong signal observed for hybridization to the human DNA control lanes, hybridization to the hybrid DNA lanes was considerably weaker, due to the lower copy number of human DNA per total

Figure 20 - Hybridization of a λHtM6 ($tRNA^{Tyr}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel.

The top panel contains DNA isolated from mouse/human somatic cell hybrid clones 1-17 in lanes A to Q. The bottom panel contains DNA from somatic cell hybrid clones 18-34 in lanes A to Q. Lane R contains mouse genomic DNA. Lane S contains human genomic DNA. All DNA samples were digested with *Bam*HI. Hybridization of a ~600 bp *Eco*RI/*Sph*I probe, encompassing a single human $tRNA^{Tyr}$ gene plus flanking sequence isolated from λHtM6 (contains a cluster of 4 human $tRNA^{Tyr}$ genes), was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was prepared by random primer extension. Membranes were washed twice at ~65° to 68°C in 0.2x SSPE, 0.1% SDS washing solution and were then exposed to X-ray film for 5 days in the presence of a Lightning Plus intensifying screen at -80°C.


A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

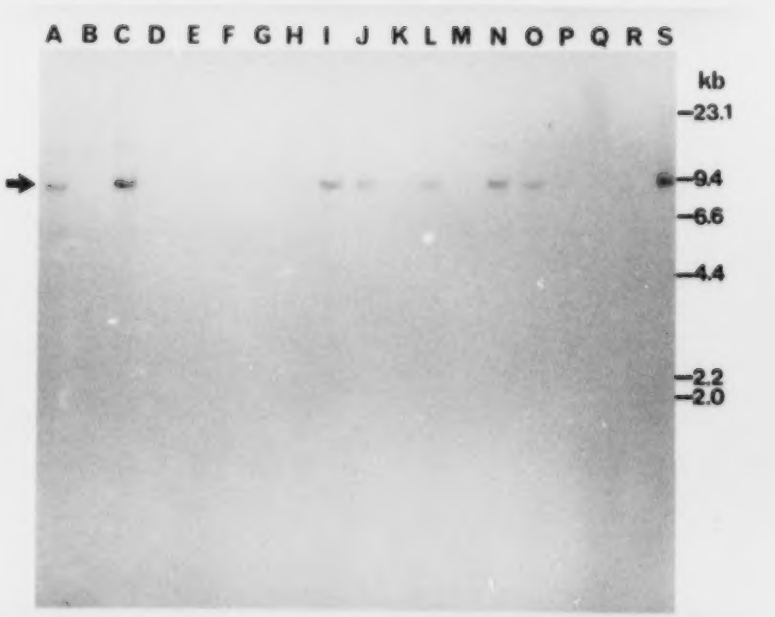
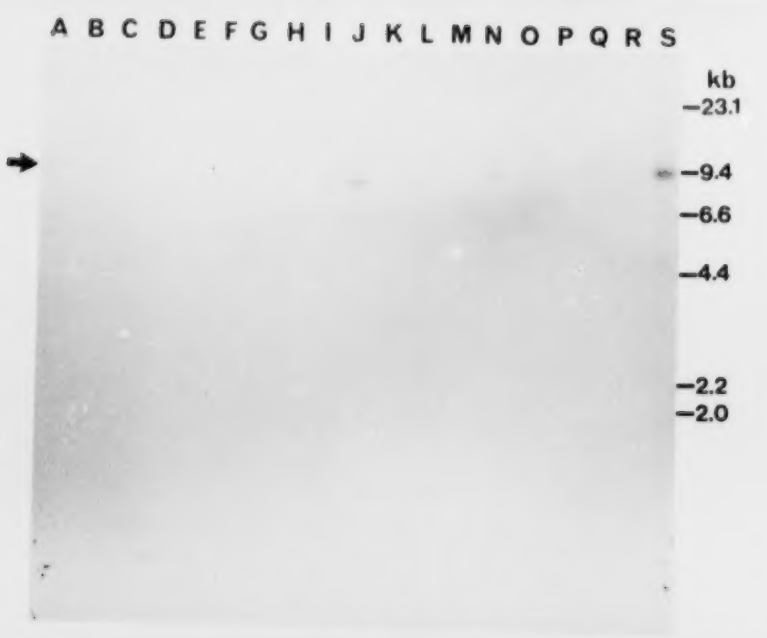

A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

Figure 21- Hybridization of a λHtM2 ($tRNA^{Tyr}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel.

The top panel contains DNA isolated from mouse/human somatic cell hybrid clones 1-17 in lanes A to Q. The bottom panel contains DNA from somatic cell hybrid clones 18-34 in lanes A to Q. Lane R contains mouse genomic DNA. Lane S contains human genomic DNA. All DNA samples were digested with *Bam*HI. Hybridization of a ~500 bp *Sty*I probe, encompassing a single human $tRNA^{Tyr}$ gene plus flanking sequence isolated from λHtM2 (contains a single human $tRNA^{Tyr}$ gene), was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. Membranes were washed twice at ~65° to 68°C in 0.2x SSPE, 0.1% SDS washing solution and were then exposed to X-ray film for 5 days in the presence of a Lightning Plus intensifying screen at -80°C.


A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

Figure 22 - Hybridization of a λHt7 ($tRNA^{Phe}$, $tRNA^{Lys}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel. The top panel contains DNA isolated from mouse/human somatic cell hybrid clones 1-17 in lanes A to Q. The bottom panel contains DNA from somatic cell hybrid clones 18-34 in lanes A to Q. Lane R contains mouse genomic DNA. Lane S contains human genomic DNA. All DNA samples were digested with *Bam*HI. Hybridization of a ~300 bp probe, encompassing a single human $tRNA^{Phe}$ gene plus flanking sequence isolated from λHt7 (contains a cluster of 2 $tRNA^{Phe}$ genes and 2 $tRNA^{Lys}$ genes) was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. Membranes were washed twice at ~65° to 68°C in 0.2x SSPE, 0.1% SDS washing solution and were then exposed to X-ray film for 5 days in the presence of a Lightning Plus intensifying screen at -80°C.

A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

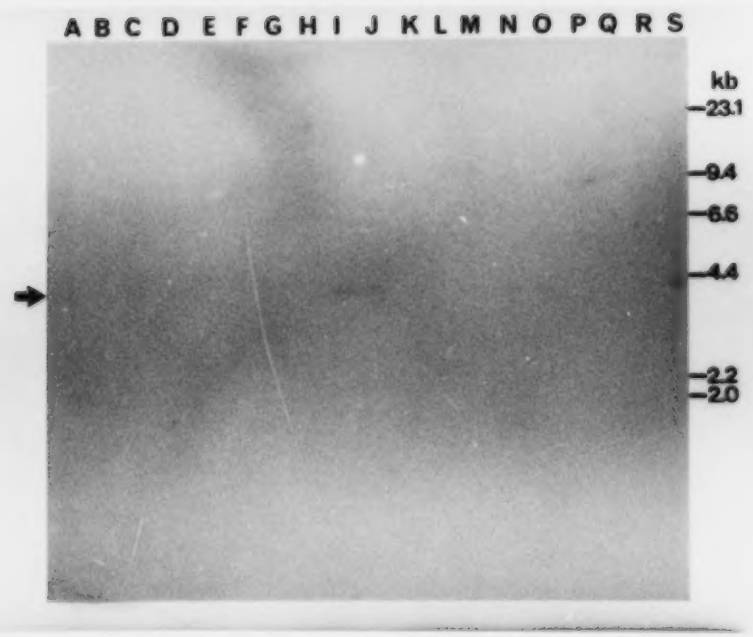
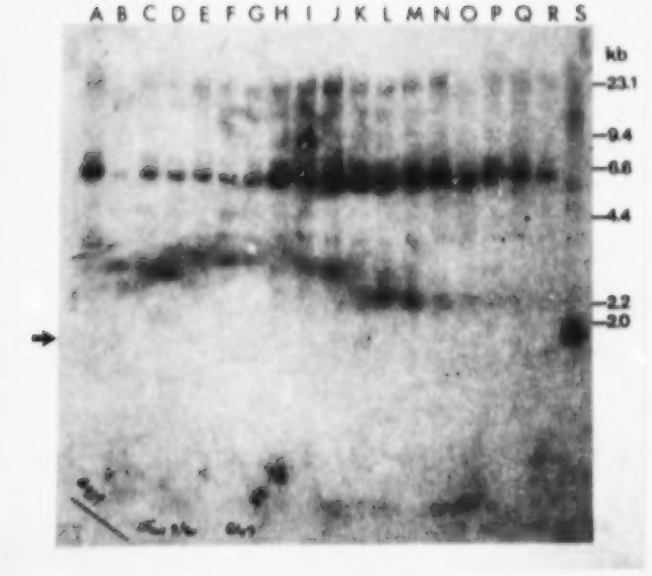
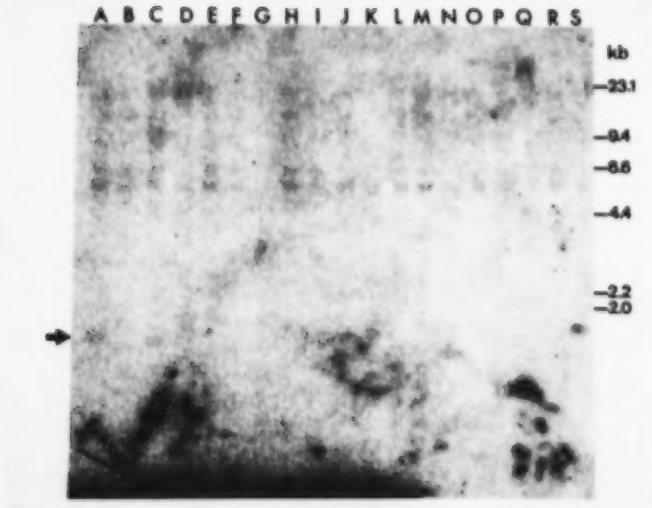

A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

Figure 23 - Hybridization of a λHt9 ($tRNA^{Gly}$) specific probe to a mouse/human somatic cell hybrid chromosomal mapping panel.

The top panel contains DNA isolated from mouse/human somatic cell hybrid clones 1-17 in lanes A to Q. The bottom panel contains DNA from somatic cell hybrid clones 18-34 in lanes A to Q. Lane R contains rodent genomic DNA. Lane S contains human genomic DNA. All DNA samples were digested with *Bam*HI. Hybridization of a ~300 bp probe, encompassing a single human $tRNA^{Gly}$ gene plus flanking sequence isolated from λHt9 (contains a single human $tRNA^{Gly}$ gene), was performed at 68°C in a modified hybridization solution containing 10x Denhardt's solution, 6x SSPE, and 7% SDS (w/v). Probe DNA was labelled by random primer extension. Membranes were washed twice at ~65° to 68°C in 0.2x SSPE, 0.1% SDS washing solution and were then exposed to X-ray film for 5 days in the presence of a Lightning Plus intensifying screen at -80°C.


A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0


A B C D E F G H I J K L M N O P Q R S
kb
23.1
9.4
6.6
4.4
2.2
2.0

mass of DNA loaded in each lane. A larger ~14 kb fragment was detected in all of the somatic cell hybrid DNA lanes plus the mouse DNA control lane, but was not observed in the human DNA control lane. The weaker signal seen for this fragment in the mouse control lane compared to the somatic cell hybrid DNA lanes probably means that less mouse genomic DNA was loaded in the control lane. Detection of hybridization of the λHtM2 specific probe ($tRNA^{Tyr}$) was complicated by the presence of significant background hybridization in one panel and weak hybridization in the other. This probe, however, demonstrated positive hybridization to a human specific band of ~9.4 kb in length. Positive hybridization was observed in the same lanes as for λHtM6, thus these two fragments must be located on the same chromosome. In those examples that exhibited weak hybridization signal, the determination of positively hybridizing lanes was done by placing the X-ray film on a light box and observing the film from an angle. The λHt7 specific probe ($tRNA^{Phe}$/$tRNA^{Lys}$) hybridized to a fragment of ~4.4 kb in length, whereas the λHt9 specific probe ($tRNA^{Gly}$) hybridized to a fragment of ~2.0 kb in length. Each of the probes exhibited weak hybridization to sequences contained within the mouse genome as determined from the hybridization observed for the mouse DNA control lane on each panel. The human chromosome complement found in each mouse-human somatic cell hybrid used in the chromosomal mapping panel and the results of hybridization of these probes to the panel are summarized in Table 1. From these data it was possible to determine the percentage of concordant or discordant hybridization of each mouse/human hybrid to each tRNA probe (see Table 2). From the results summarized in Table 2 it can be seen that the human DNA contained within the human-λ recombinants λHtM2 and λHtM6 are localized on chromosome 14. The human DNA from λHt7 is located on chromosome 11, whereas the human DNA from λHt9 is found on chromosome 1. More precise determination of chromosomal locations was not possible using this particular mapping panel. A summary of known chromosomal locations of human tRNA genes is shown in Table 3.

Table 1 - Summary for hybridization of tRNA gene-containing probes to a human chromosomal mapping panel.

This table shows the human chromosome complement contained within each mouse/human somatic cell hybrid (+ = presence of chromosome: - = absence of chromosome: Δ = very low proportion of that human chromosome detected in metaphase spreads from that somatic cell hybrid clone) that was used to provide DNA for the chromosomal mapping panel and a summary of the results of hybridization for each tRNA gene-containing probe to DNA isolated from each somatic cell hybrid (+ = positive hybridization: - = negative hybridization)

Clone	Hybridization				Chromosome																							Translocations
	HJ17	HJ19	HJM2	HJM6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
18	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-	-	+	Δ	5/X
1	-	-	+	+	-	+	-	-	-	-	+	+	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	
2	+	-	+	+	-	-	+	-	+	-	-	-	-	-	+	-	-	+	-	-	+	+	-	-	+	-	-	
3	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-	-	+	-	-	+	+	-	-	-	+	
20	+	+	+	+	+	+	+	-	+	+	+	-	-	+	+	+	-	+	Δ	+	+	+	+	+	+	+	Δ	X/15, 15/X
19	-	-	-	-	-	-	+	-	+	-	-	+	-	+	-	+	+	-	+	+	-	-	-	+	-	-	+	
23	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	+	-	+	
24	+	+	-	-	+	+	+	-	+	-	Δ	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	7/9
22	-	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	
26	+	-	+	+	Δ	Δ	+	+	-	+	-	+	-	+	+	+	+	+	+	-	+	+	-	+	+	-	-	2/1
27	+	-	+	+	Δ	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	-	+	+	-	+	1/2
32	+	-	+	+	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+	-	-	-	-	-	-	+	
7	-	-	+	+	-	-	+	+	+	-	+	+	Δ	+	-	+	-	+	+	+	+	+	-	+	+	-	-	17/9
6	-	-	+	+	-	-	-	-	+	-	-	+	Δ	+	-	+	+	+	+	+	+	+	-	+	+	+	-	9/17
21	+	-	-	-	-	-	-	+	-	-	+	-	-	-	+	+	+	-	-	+	-	-	-	+	+	+	+	
29	-	-	+	+	-	-	+	-	-	-	-	-	-	+	-	-	-	+	+	-	-	+	-	-	-	-	-	
30	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	Δ	Δ	22/X
28	+	-	-	-	+	+	+	+	-	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	Δ	Δ	X/22
31	+	-	+	+	-	-	-	+	-	-	+	-	-	+	+	-	-	+	-	+	+	-	-	-	+	-	+	
25	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	+	+	+	
4	+	-	-	-	-	+	+	+	-	-	-	-	-	+	+	-	+	-	-	+	+	+	-	+	+	-	-	
5	-	-	-	-	-	+	Δ	-	+	+	-	+	-	+	-	+	-	+	+	-	Δ	+	-	+	+	-	+	17/3
14	+	-	+	+	-	-	-	-	+	-	+	-	-	+	+	-	+	+	-	-	+	-	-	+	+	-	-	
15	+	-	-	-	-	+	-	-	-	+	+	+	-	+	+	-	-	-	+	-	+	-	+	+	+	+	-	
11	-	+	+	+	+	-	+	-	+	-	+	+	-	+	-	+	-	+	+	-	+	-	-	-	-	-	+	
12	+	-	+	+	-	+	+	+	-	+	+	-	-	+	+	+	+	+	+	-	+	+	-	+	+	-	+	
8	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	-	-	+	-	+	-	-	-	+	-	+	
9	+	-	+	+	-	+	-	+	+	+	+	+	-	+	+	-	-	+	-	-	+	-	+	+	+	-	+	
10	+	-	+	+	-	-	+	+	+	-	+	+	-	+	+	+	-	+	-	-	+	+	+	+	+	-	+	
16	+	-	-	-	Δ	-	-	+	+	-	+	-	-	+	+	+	-	-	-	-	+	-	-	+	-	-	+	X/1
17	+	-	+	+	-	-	+	-	-	-	Δ	+	+	+	+	+	-	+	-	-	+	+	-	+	-	-	+	ISO7p
15	+	-	+	+	Δ	-	-	-	+	+	+	-	-	+	+	+	-	+	-	-	+	-	+	+	-	+	Δ	1/X
33	-	-	-	-	-	-	Δ	-	+	-	-	+	-	+	-	+	+	+	-	-	-	+	-	+	+	-	Δ	3/X
34	-	-	-	-	-	-	Δ	-	-	-	-	-	+	Δ	-	+	-	-	-	-	-	-	-	+	+	-	Δ	3/X, 10q-

Table 2 - Percentage discordancy table for hybridization of tRNA gene-containing probes to a human chromosomal mapping panel.

The summary of hybridization results for concordant hybridization (+/+ or -/-) or discordant hybridization (+/- or -/+) is indicated for the hybridization of the λHt7 specific probe (a), λHt9 specific probe (b), λHtM2 specific probe (c), and the λHtM6 specific probe (d), to DNA for each human chromosome. Concordant hybridization is the result where hybridization is observed when a chromosome is present or hybridization is not observed when a chromosome is absent. Discordant hybridization is the result where hybridization is observed when a chromosome is absent or when hybridization is not observed when a chromosome is present. Percentage discordant hybridization is calculated by dividing the number of discordant results by the total number of observations (discordant + concordant hybridizations). A low % discordancy of hybridization for a chromosome indicates that the DNA fragment used in the probe is probably located on that chromosome. Discordant results due to absence of hybridization to DNA from somatic cell hybrid clones containing very low proportions of a given human chromosome are not included in the % discordancy calculations. Somatic cell hybrid clones 5, 22, 24, 30, 33 were not used for the determination of % discordant hybridzation for the λHtM2 and λHtM6 probes due to a low percentage of human chromosome 14 detected in metaphase spreads. Clones 24 and 28 were eliminated from the calculations because very little DNA was electrophoresed for these samples. Somatic cell hybrid clone 8 was likewise not included due to human error. Somatic cell hybrid clones 2 and 28 were not used for the determination of % discordant hybridization for the λHt9 probe due to the low percentage of human chromosome 1 detected in metaphase spreads.

a) λHt7 ($tRNA^{Phe}$/ $tRNA^{Lys}$)

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
(+/+)	3	8	10	10	9	7	12	7	4	15	18	11	7	14	6	8	16	10	6	14	14	5	9
(+/−)	11	9	8	8	9	11	4	11	14	3	0	7	11	4	11	10	2	8	12	4	4	12	6
(−/+)	2	5	8	5	9	3	6	9	2	10	0	10	8	10	9	4	7	8	2	7	8	3	8
(−/−)	14	11	5	11	7	13	10	7	12	5	16	6	8	6	7	12	8	8	14	9	8	12	4
% discordancy	43	42	52	38	53	41	31	59	50	39	0	50	56	41	61	41	27	47	41	32	35	47	52

b) λHt9 ($tRNA^{Gly}$)

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
(+/+)	4	3	4	1	4	2	3	3	2	4	2	4	2	4	3	3	4	3	2	2	2	3	1
(+/−)	0	1	0	3	0	2	0	1	2	0	2	0	2	0	0	1	0	1	2	2	2	1	1
(−/+)	0	9	12	13	13	8	14	12	3	20	14	16	13	18	11	8	17	13	5	19	18	5	16
(−/−)	24	18	13	15	15	20	13	16	23	7	14	12	15	10	17	20	10	15	23	9	10	22	8
% discordancy	0	32	41	50	41	31	47	41	17	65	50	50	47	56	35	28	55	44	22	66	63	19	65

c) λHtM2 ($tRNA^{Tyr}$)

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
(+/+)	3	6	11	8	11	7	12	9	3	15	12	11	7	18	8	7	16	11	5	11	11	4	8
(+/−)	12	11	7	10	7	11	5	9	13	3	6	7	11	0	9	11	2	7	13	7	7	14	7
(−/+)	0	4	3	5	4	2	4	2	1	4	4	5	5	0	3	3	4	2	2	7	6	3	6
(−/−)	8	5	5	4	5	7	5	7	8	4	5	4	4	9	6	6	5	7	7	2	3	6	2
% discordancy	52	58	38	56	41	48	35	41	56	27	37	44	59	0	46	52	22	33	56	52	48	63	57

d) λHtM6 ($tRNA^{Tyr}$)

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
(+/+)	3	6	11	8	11	7	12	9	3	15	12	11	7	18	8	7	16	11	5	11	11	4	8
(+/−)	12	11	7	10	7	11	5	9	13	3	6	7	11	0	9	11	2	7	13	7	7	14	7
(−/+)	0	4	3	5	4	2	4	2	1	4	4	5	5	0	3	3	4	2	2	7	6	3	6
(−/−)	8	5	5	4	5	7	5	7	8	4	5	4	4	9	6	6	5	7	7	2	3	6	2
% discordancy	52	58	38	56	41	48	35	41	56	27	37	44	59	0	46	52	22	33	56	52	48	63	57

Table 3 - Summary of known chromosomal locations for human tRNA gene loci. This table is a compilation of data reported in this study and those reported in McAlpine et al., 1990. An (*) indicates tRNA gene locations determined in this report. The designation for these genes is tentatively proposed and is not officially recognized by the HGM nomenclature committee.

Designation	tRNA Gene Description	Chromosomal Location
TRE	glutamic acid	1p36
TREL1	glutamate acid-like	17pter-p12
TRF1*	phenylalanine 1	11
TRF2*	phenylalanine 2	11
TRG1	glycine 1 (GCC)	16
TRGL1	glycine (CCC)-like1	1pter-p34
TRGP1	glycine (GCC) pseudogene1	1pter-p34
TRG2*	glycine 2	1
TRK1	lysine 1	17pter-p12
TRK2*	lysine 2	11
TRK3*	lysine 3	11
TRL1	leucine 1	14q11-q12
TRL2	leucine 2	17pter-p12
TRM1	methionine (initiator) 1	6p23-q12
TRM2	methionine (initiator) 2	6p23-q12
TRN	asparagine	1p36
TRP1	proline 1	14q11-q12
TRP2	proline 2	14q11-q12
TRP3	proline 3 (UGG)	5
TRQ1	glutamine 1	17pter-p12
TRSP	phosphoserine (opal suppressor)	19
TRSPP1	phosphoserine (opal suppressor) pseudogene 1	22q11
TRT2	threonine (UGU) 2	14q11-q12
TRV	valine	5
TRY1*	tyrosine 1	14
TRY2*	tyrosine 2	14
TRY3*	tyrosine 3	14
TRY4*	tyrosine 4	14

3.3 Isolation of human tRNA genes

3.3.1 tRNA PCR

A. tRNA^{Tyr}: The results of PCR using primers specific for the 5'- and 3'-conserved regions of human tyrosine tRNA genes are shown in Figure 24. Amplification products were size fractionated on a 1.5% agarose gel and were transferred to nylon membrane. Immobilized DNA was screened by hybridization to the 3' gene-specific primer. When only the 5' gene-specific primer was used individually in a PCR, at least 16 faint bands were visualized by ethidium bromide staining of the agarose gel. None of these was shown to hybridize to the probe sequence corresponding to the 3'- end of the gene, thus these amplification products are devoid of the 3' tRNA^{Tyr} gene sequences. When the 3' gene-specific primer was used individually in a PCR, no amplification products were detectable, either by ethidium bromide staining or by hybridization with this primer. When both primers were utilized in a PCR, strongly hybridizing fragments of ~0.67, 0.63 and 0.1 kb in length were observed. The smaller (~0.1 kb) fragments are tRNA^{Tyr} gene plus intron sequences. Amplification products from this PCR were cloned into the *Hinc*II site of pUC118 for sequence analysis.

The larger hybridizing bands from the PCR involving both gene primers were cloned and sequenced. A schematic representation of these fragments is shown in Figure 25. A comparison of the sequences of 5 fragments is shown in Figure 26. The fragments ranged in size from 652 bp to 719 bp in length. The major sequence differences between these 5 fragments are found largely in the intron sequences. The intergenic flanking sequences were shown to be nearly identical between all the fragments characterized. These fragments were shown to have an overall sequence homology of at least 95%. When the intron sequences are omitted, for the purposes of sequence comparisons, the fragments exhibit at least 98% homology.

Figure 24 - tRNA PCR for the human $tRNA^{Tyr}$ gene family.

(*Left*) An ethidium bromide stained agarose gel (1.5%) showing the results of electrophoretic fractionation of products generated by single primer PCR using only the 5' gene-specific (lane A), or using only the 3' gene-specific primer (lane B). Lane C shows the results for PCR using both the 5' and 3' the gene-specific primers. (*Right*) DNA from each PCR was transferred to nylon membrane and hybridized with the $tRNA^{Tyr}$ 3' gene-specific primer.

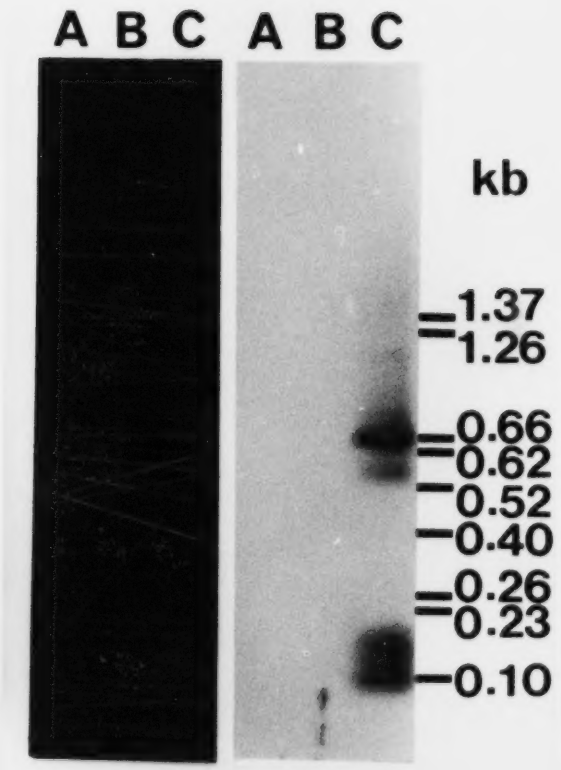

A B C A B C
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

Figure 25 -Schematic representation of the amplification products from tRNA PCR using primers specific for the human $tRNA^{Tyr}$ gene family. Rectangles represent either a complete tRNA gene or part of a tRNA gene. The 5' terminus of a tRNA gene is represented as a darkened block. Triangles represent the 3' terminus of a tRNA gene. The shaded regions represent intron sequences. The designation "pJM4" indicates intron sequences previously characterized by MacPherson (1988). The designation "Vold" indicates the intron sequence previously characterized by Green et al., (1990). Other designations (e.g. "6.5" - intron from 5' tRNA gene contained on pHy6) indicate intron sequences which have not previously been characterized. Partial tRNA gene sequences are shown as smaller rectangles containing a darkened block and/or triangle figure. Small arrows indicate the primers that were used in the PCRs which generated these fragments (LTyr and RTyr - see Figure 8). The intron specific primers (Tyr3, Tyr4, Tyr5, and Tyr6) are unique for the different introns found within the $tRNA^{Tyr}$ genes. The flanking sequence specific primers (Tyr7 and Tyr8) were used in a PCR with the the intron specific primers using, as template, either plasmid DNA containing the complementary intron sequence or human genomic DNA. The oligonucleotide Tyr9 (solid bar) was used to hybridize to amplification products from these latter PCRs.

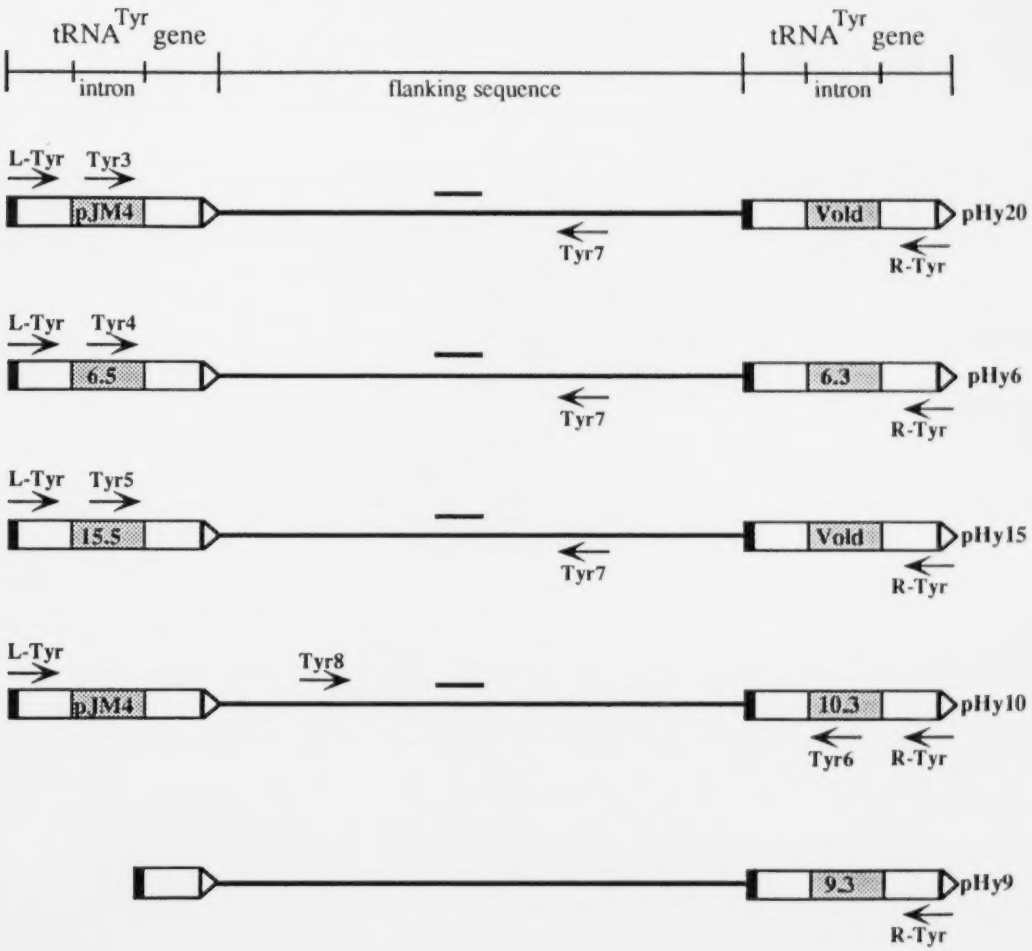
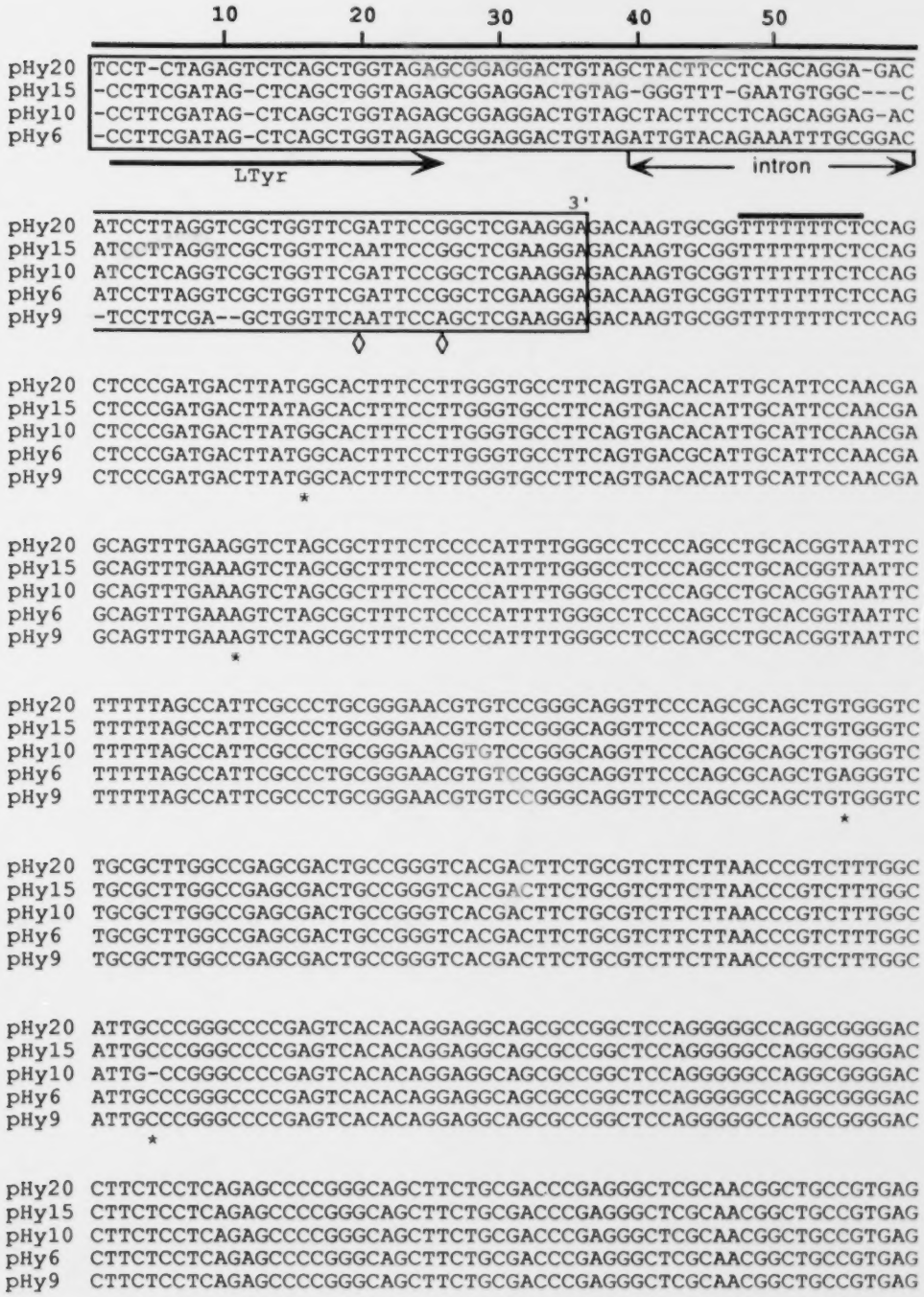

tRNA^Tyr gene
tRNA^Tyr gene
intron
flanking sequence
intron
L-Tyr
Tyr3
pJM4
Vold
pHy20
Tyr7
R-Tyr
L-Tyr
Tyr4
6.5
6.3
pHy6
Tyr7
R-Tyr
L-Tyr
Tyr5
15.5
Vold
pHy15
Tyr7
R-Tyr
L-Tyr
Tyr8
pJM4
10.3
pHy10
Tyr6
R-Tyr
9.3
pHy9
R-Tyr

Figure 26 - Nucleotide sequence comparisons of $tRNA^{Tyr}$ gene-containing recombinants pHy6, pHy9, pHy10, pHy15 and pHy20. Boxed regions indicate $tRNA^{Tyr}$ gene sequences. The solid line above the sequence downstream from the 3' end of the first gene indicates a putative RNA polymerase III termination signal. The primers (LTyr and RTyr - see Figure 8) used in the PCR are indicated by thick arrows. Intron sequences are shown. Asterisks (*) indicate sequence differences in the flanking (i.e. intergenic) regions of the five fragments characterized. Diamonds (◊) indicate previously characterized tRNA gene sequence polymorphisms.


10
20
30
40
50
LTyr
intron
3'

```
pHy20  GAGGTGGGGGGTCCGCGGGAAGAGGTATCTGGCGCTCCCGGGACCTGGGAATCAGAAAGA
pHy15  GAGGTGGGGGGTCCGCGGGAAGAGGTATCTGGCGCTCCCGGAACCTGGGAATCAGAAAGA
pHy10  GAGGTGGGGGGTCCGCGGGAAGAGGTATCTGGCGCTCCCGGAACCTGGGAATCAGAAAGA
pHy6   GAGGTGGGGGGTCCGCGGGAAGAGGTATCTGGCGCTCCCGGAACCTGGGAATCAGAAAGA
pHy9   GAGGTGGGGGGTCCGCGGGAAGAGGTATCTGGCGCTCCCGGAACCTGGGAATCAGAAAGA
                                                  *

pHy20  GAGAACACAATACTAAAAACACGAAGCCTAAAAATGACACAATGTTATGGAGACAAGGCG
pHy15  GAGAACACAATACTAAAAACACGAAGCCTAAAAATGACACAATGTTATGGAGACAAGGCG
pHy10  GAGAACACAATACTAAAAACACGAAGCCTAAAAATGACACAATGTTATGGAGACAAGGCG
pHy6   GAGAACACAATACTAAAAACACGAAGCCTAAAAATGACACAATGTTATGGAGACAAGGCG
pHy9   GAGAACACAATACTAAAAACACGAAGCCTAAAAATGACACAATGTTATGGAGACAAGGCG

                                                5'
pHy20  GCACCCGGGAAGCTGTGCCCGCTCCCTTCGATAGCTCAGCTGGTAGAGCGGAGGACTGTA
pHy15  GCACCCGGGAAGCTGTGCCCGCTCCCTTCGATAGCTCAGCTGGTAGAGCGGAGGACTGTA
pHy10  GCACCCGGGAAGCTGTGCCCGCTCCCTTCGATAGCTCAGCTGGTAGAGCGGAGGACTGTA
pHy6   GCACCCGGGAAGCTGTGCCCGCTCCCTTCGATAGCTCAGCTGGTAGAGCGGAGGACTGTA
pHy9   GCACCCGGGAAGCTGTGCCCGCTCCCTTCGATAGCTCAGCTGGTAGAGCGGAGGACTGTA

pHy20  G-GCGC-GCGCCCGT-GGC--CATCCTTAGGTCGCTGGTTCGATTCCGGCTCGAAGGA
pHy15  G-GCGC-GCGCCCGT-GGC--CATCCTTAGGTCGCTGGTTCAATTCCAGCTCGAAGGA
pHy10  G-CTACTTCCTCAGCAGGAGACATCCTTAGGTCGCTGGTTCAATTCCGGCTCGAAGGA
pHy6   GTTGGCTGTGTCCTT-AGA--CATCCTTAGGTCGCTGGTTCGATTCCGGCTCGAAGGA
pHy9   G-GCGC-GCGCCCGC-GGC--CATCCTTAGGTCGCTGGTTCGATTCCGGCTCGAAGGA
```

Four of the five fragments contain complete $tRNA^{Tyr}$ genes positioned at both termini of the fragments. The fifth fragment (pHy9) exhibits a $tRNA^{Tyr}$ gene at one terminus while the other terminus contains a sequence that appears to be the complete 3' gene region plus a hybrid of the sequences found within the primer sequences. All genes found on these fragments were shown to be arranged as direct tandem repeats separated by 528 bp of intergenic flanking sequence. Despite the extremely high sequence similarities, there were a number of examples where the sequences differed from each other. The intergenic sequence of all of these examples was shown to very similar to some other $tRNA^{Tyr}$ gene-containing fragments previously characterized. One clone, pHy20, was shown to be 98% homologous to the sequence found on pJM4, including the conservation of the intron sequences and the relative arrangement of two $tRNA^{Tyr}$ genes as direct tandem repeats (MacPherson, 1988; D. Spadafora, unpublished).

To determine if the other large fragments containing different $tRNA^{Tyr}$ genes at the termini were artifacts, PCR was performed using primers that were uniquely specific for the different intron sequences in combination with a second primer that was specific for the common intergenic flanking regions. PCR was performed using each set of primers on the plasmid DNA containing that particular intron sequence and also, separately on human genomic DNA. The results are shown in Figure 27. Amplification from both human genomic DNA and from plasmid DNA was observed for the PCRs using primers specific for the introns contained in pHy20 and pHy6. Amplification for those PCRs using primers specific for the introns contained in pHy10 and pHy15 was seen only for the plasmid DNA but not from human genomic DNA. All amplified fragments hybridized to a probe that was specific for a region that is contained between the primers used for PCR. These results indicate that the sequences contained in pHy6 and pHy20 are likely to be contained in the human genome, while those sequences from pHy10 and pHy15 are artifacts.

Figure 27 - Amplification product from PCR using various $tRNA^{Tyr}$ intron-specific primers and a primer common to the intergenic flanking sequence found within pHy6, pHy10, pHy15 and pHy20.

Lanes A and B: PCR was performed using a primer specific for the intron contained within the upstream $tRNA^{Tyr}$ gene found on pHy20 (Tyr3) and a second primer (Tyr 7) that was specific for the flanking sequence, using human genomic DNA (lane A) or pHy20 DNA (lane B) as the template .

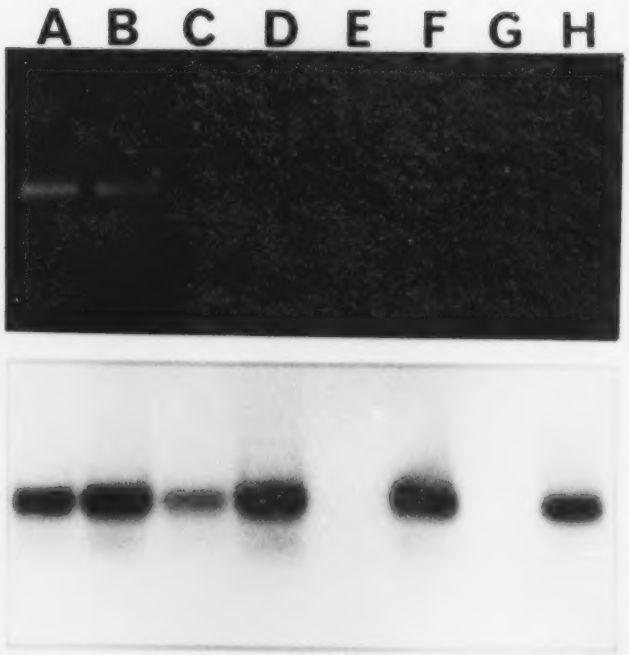
Lanes C and D: PCR was performed using a primer specific for the intron contained within the upstream $tRNA^{Tyr}$ gene found on pHy6 (Tyr4) and a second primer (Tyr 7) that was specific for the flanking sequence, using human genomic DNA (lane C) or pHy6 DNA (lane D) as the template.

Lanes E and F: PCR was performed using a primer specific for the intron contained within the upstream $tRNA^{Tyr}$ gene found on pHy15 (Tyr5) and a second primer (Tyr 7) that was specific for the flanking sequence, using human genomic DNA (lane E) or pHy15 DNA (lane F) as the template.

Lanes G and H: PCR was performed using a primer specific for the intron contained within the downstream $tRNA^{Tyr}$ gene found on pHy10 (Tyr6) and a second primer (Tyr 8) that was specific for the flanking sequence, using human genomic DNA (lane G) or pHy10 DNA (lane H) as the template.

A schematic representation of primers used is shown in Figure 26.

(*Top*) Amplification products were size fractionated on a 0.75% agarose gel and stained with ethidium bromide. (*Bottom*) DNA was transferred to nylon membrane and hybridized to an oligonucleotide probe (Tyr9) which should be contained within each of the amplification products if amplification occurred from the sequences contained within each of the plasmid clones characterized.


A B C D E F G H

A comparison of the intron sequences for the tyrosine tRNA genes is shown in Figure 28. From this, it can be seen that the introns for these genes are either 16, 18, 20 or 21 bp in length. Many of these intron sequences differ by only a single base pairfrom other examples while some examples are significantly different. There appears to be conservation of some bases in the intron sequence, most notably the nucleotides adjacent to the 3' splice sites between intron and exon sequences. At least fourteen different intron sequences for $tRNA^{Tyr}$ genes and pseudogenes have been characterized.

B. $tRNA^{Phe}$: Amplification products resulting from PCR using $tRNA^{Phe}$ gene-specific primers were separated on 1.5% agarose gels, the DNA transferred to nylon membrane and the sequences probed with the 3' gene-specific primer. Both the 5' gene-specific and 3' gene-specific primers were utilized in PCRs individually or with each other. The results of PCR using these primers are shown in Figure 29. When the 5' gene-specific primer was used individually, bands of ~0.8 and 1.2 kb in length were visualized by staining with ethidium bromide. Of these, only the 1.2 kb fragment hybridized when probed with the 3' gene-specific primer. This indicated that this fragment contained at least both termini of a $tRNA^{Phe}$ gene. Two different bands of ~450 and 500 bp in length were produced when the 3' gene-specific primer was used individually in a PCR. Both of these bands hybridized to the 3' gene-specific primer, as expected. When both of the primers were used in a PCR, an intense band was produced of ~90 bp in length that would be expected to be strictly the tRNA gene coding sequence (~72 bp). Unexpectedly, a band of ~ 500 bp that hybridized to the 3' gene-specific primer was also amplified.

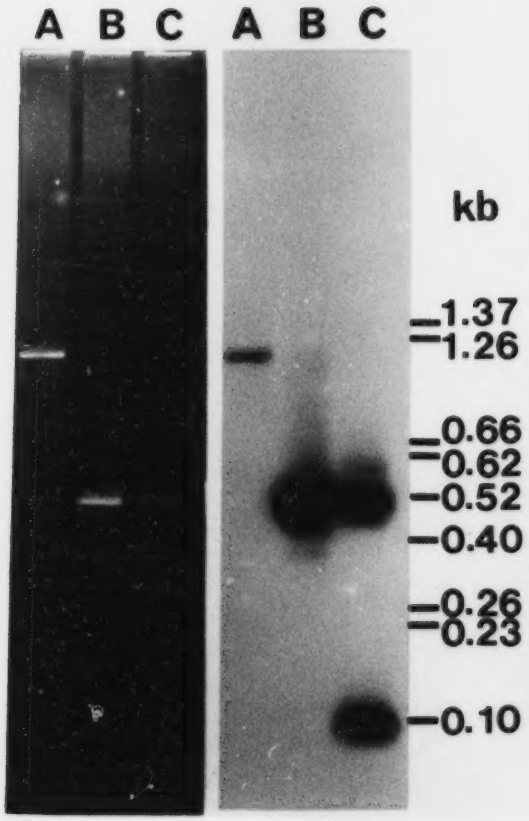
Figure 28 - Nucleotide sequence comparison of $tRNA^{Tyr}$ gene intron sequences.

Introns from a *Xenopus laevis* $tRNA^{Tyr}$ gene and from human $tRNA^{Tyr}$ genes isolated by tRNA PCR or previously isolated from a human-λ recombinant library are shown. The intron sequences for pJM4, pM6, pM612, pJM6128, pJM6IT were previously determined by MacPherson (1988). The intron sequence for pHtT2 is from van Tol and Beier (1988). pHtT2 is a probably a $tRNA^{Tyr}$ pseudogene. The sequence for pHy20.3 is the same as that determined by Green et al., (1990). The remaining human sequences are novel to this study. The intron designations pHy10.3, pHy6.3, pHy15.3, pHy9.3 and pHy20.3 refer to the introns contained within the downstream $tRNA^{Tyr}$ genes of the respective fragments represented in Figure 27. The intron pHy6.5 is contained in the upstream $tRNA^{Tyr}$ gene found in pHy6. The introns for pHy8 and pHy17 were found only on solitary $tRNA^{Tyr}$ gene fragments and were not found on any of the larger fragments containing tandemly repeated $tRNA^{Tyr}$ genes. Asterisks (*) indicate highly conserved nucleotides.

I.D.	INTRON SEQUENCE
X. LAEVIS	-GTGTGA-TCGAG-CA
pJM4	CTACTTCCTC-AG-CA-----GGAGAC
pHy10.3	CTACTTCCTCCAG-CA-----GGAGAC
pJM6128	ATTGT-AC---AGACA--TTTGCGGAC
pM612	ATTGT-A-T--AGACA--TTTGCGGAC
pHy6.5	ATTGT-AC---AGAAA--TTTGCGGAC
pHy17	ACTGCG------GAAACGTTTGAGGAC
pM6	ACTGCG------GAAACGTTTGTGGAC
pM6IT	CCTGT-A-----GAAACATTTGTGGAC
pHy8	CC---GA-T---GAAACATTTGTGGAC
pHy6.3	-TTG-G-CT-GTG-----TCCTTAGAC
pHy15.3	---GGGTTT---GAA-----TGTGGTC
pHy9.3	---GCGCGC---G-----CCCGCGGCC
pHy20.3	---GCGCGC---G-----CCCGTGGCC
pHtT2	---GTGCAC---G-----CCCGTGGCC
	* ****

Figure 29 - tRNA PCR using primers specific for the human $tRNA^{Phe}$ gene family.

(*Left*) An ethidium bromide stained 1.5% agarose gel showing the amplification products for single primer PCR using only the 5' gene-specific primer (lane A), or the 3' gene-specific primer (lane B). Lane C shows the results of amplification for PCR using both primers. (*Right*) The amplification products were transferred to nylon membrane and hybridized with the $tRNA^{Phe}$ 3' gene-specific primer.


A B C A B C
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

The amplification products from each of the reactions were then cloned into the *Hinc*II site of pUC118 for sequence analysis. A schematic representation of the results is shown in Figure 30. The sequence results are shown in Figures 31 and 32. The hybridizing fragment from PCR using the 5' gene-specific primer (designated pHf1-39) was shown to be 1152 bp in length. Sequence comparisons revealed no sequence similarities to other previously characterized $tRNA^{Phe}$ flanking sequences. This fragment contained a complete tRNA located at one fragment end while the other end was shown to be a partial $tRNA^{Phe}$-like sequence that lacks the 3' region of the gene. The tRNA gene and tRNA gene-like sequence were found arranged as inverted repeats relative to each other, separated by 1057 bp. A potential terminator sequence was found 6 nt downstream from the 3' end of the gene.

The PCR products from the 3' gene-specific primer PCR were also sequenced. These fragments were shown to be 491 bp and 426 bp in length. The longer fragment (designated pHf7-31) contains a complete $tRNA^{Phe}$ gene and the 3' region of a $tRNA^{Phe}$ gene positioned at its termini. Within this fragment, a repeated sequence with a consensus repeat sequence consisting of $(GAXT)_{13}$ was found downstream from one of the primer binding sites. The shorter fragment (designated pHf7-30) was the result of amplification between two sequences corresponding to the 3' region of a $tRNA^{Phe}$ gene. Examination of this sequence revealed a potential pseudogene of 72 bp in length, with an eight bp intron sequence. The actual sequence of this pseudogene (outside the primer binding site) is unlike any tRNA gene sequence previously characterized, however, it is potentially capable of forming a cloverleaf structure. A number of direct repeats, of unknown significance, were also observed. The amplification products from PCR using both gene-specific primers consisted of $tRNA^{Phe}$ gene sequences and also one of the PCR products characterized for the PCR using the 3' gene-specific primer alone (pHf7-31). The larger amplification products found in the PCR for the 5' gene-specific primer condition were not found.

Figure 30 - Schematic representation of amplification products from tRNA PCR using primers specific for the human $tRNA^{Phe}$ gene family. Rectangles represent either a complete tRNA gene or part of a tRNA gene. The darkened block within each box represents the 5' terminus of a tRNA gene while the arrowheads represent the 3' terminus of a tRNA gene. Partial tRNA gene sequences are represented as boxes containing only the darkened block or arrowhead figure. Solid lines represent flanking sequence. The primers (RPhe or LPhe) that were used in a PCR to generate each fragment are shown as arrows.

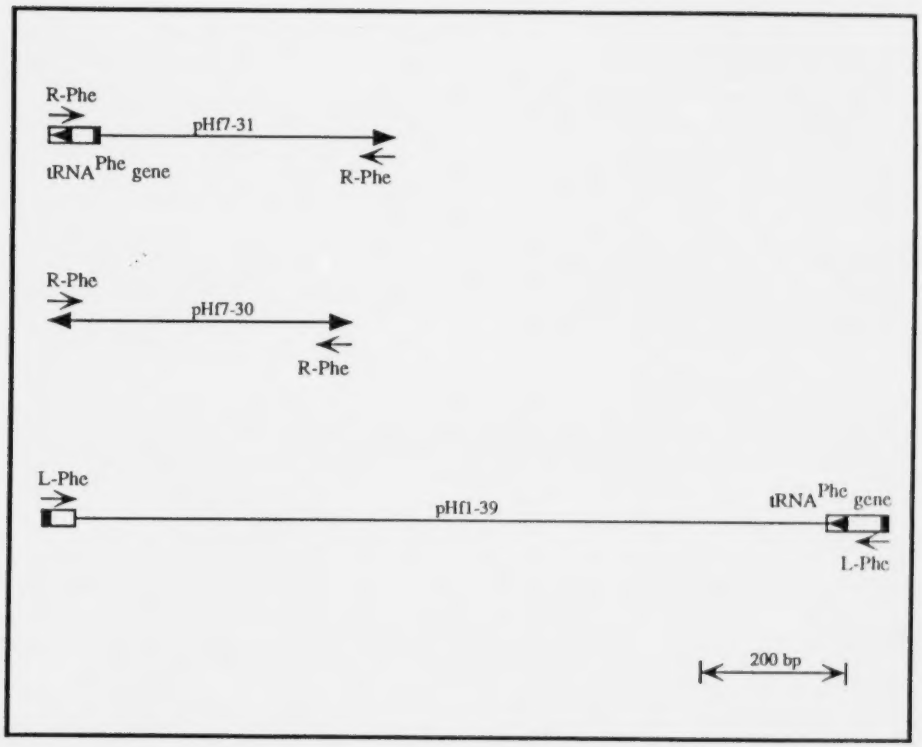
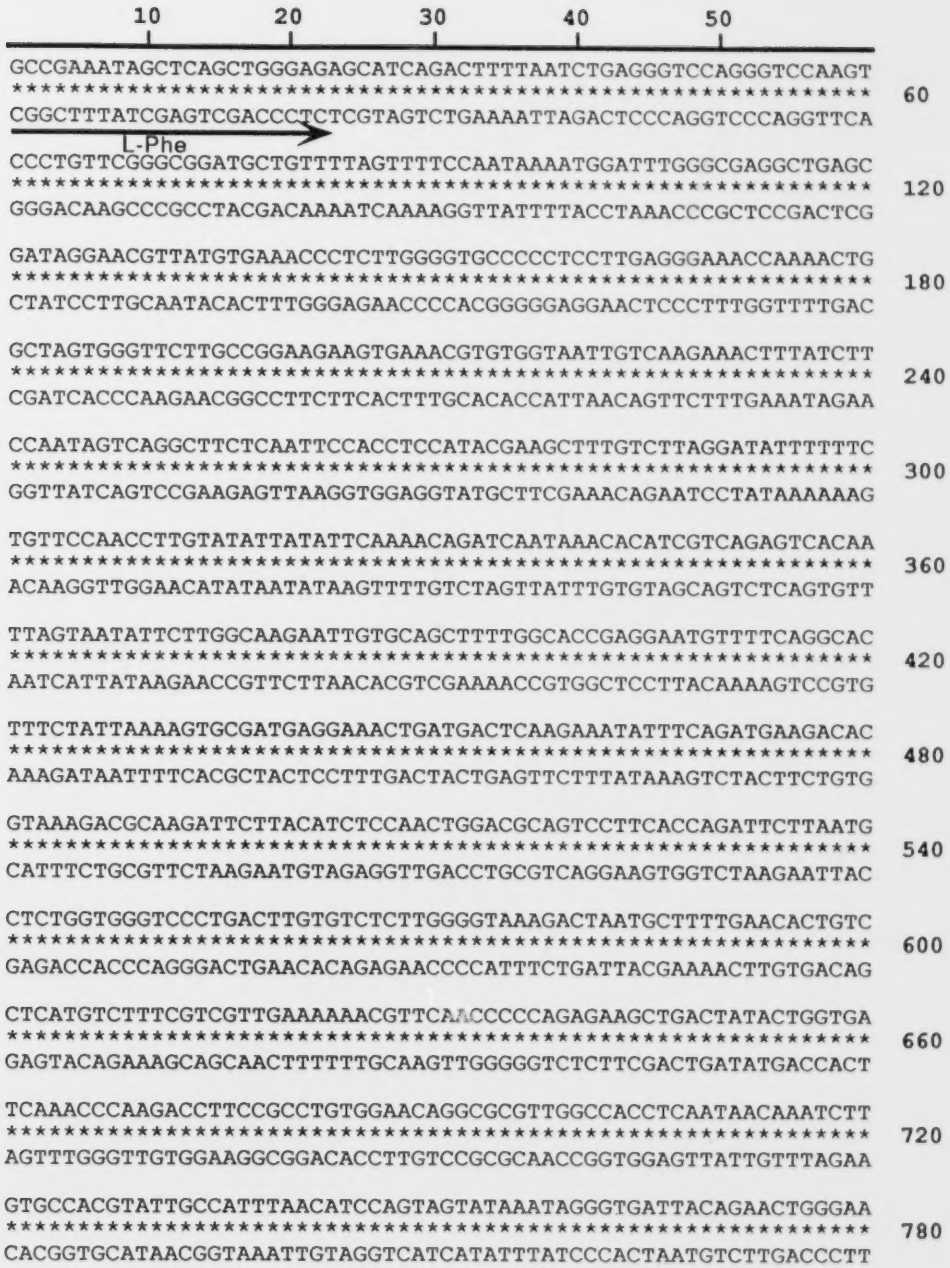

R-Phe
pHf7-31
tRNA Phe gene
R-Phe
R-Phe
pHf7-30
R-Phe
L-Phe
pHf1-39
tRNA Phe gene
L-Phe
200 bp

Figure 31 - Nucleotide sequence of a human $tRNA^{Phe}$ gene and flanking sequence contained within pHf1-39.

The human DNA fragment within this recombinant was obtained by tRNA PCR using only the $tRNA^{Phe}$ 5' gene-specific primer. One terminus of this recombinant contains the complete $tRNA^{Phe}$ gene coding sequence while the other terminus is derived from a region containing only the sequence corresponding to the 5' gene-specific primer. Thick arrows indicate the primer binding sites. A solid bar indicates a putative RNA polymerase III signal in the 3' flanking region of the complete tRNA gene sequence.


10
20
30
40
50
L-Phe
GCCGAAATAGCTCAGCTGGGAGAGCATCAGACTTTTAATCTGAGGGTCCAGGGTCCAAGT
CGGCTTTATCGAGTCGACCCTCTCGTAGTCTGAAAATTAGACTCCCAGGTCCCAGGTTCA
60
CCCTGTTCGGGCGGATGCTGTTTTAGTTTTCCAATAAAATGGATTTGGGCGAGGCTGAGC
GGGACAAGCCCGCCTACGACAAAATCAAAAGGTTATTTTACCTAAACCCGCTCCGACTCG
120
GATAGGAACGTTATGTGAAACCCTCTTGGGGGTGCCCCCTCCTTGAGGGAAACCAAAACTG
CTATCCTTGCAATACACTTTGGGAGAACCCCACGGGGGAGGAACTCCCTTTGGTTTTGAC
180
GCTAGTGGGTTCTTGCCGGAAGAAGTGAAACGTGTGGTAATTGTCAAGAAACTTTATCTT
CGATCACCCAAGAACGGCCTTCTTCACTTTGCACACCATTAACAGTTCTTTGAAATAGAA
240
CCAATAGTCAGGCTTCTCAATTCCACCTCCATACGAAGCTTTGTCTTAGGATATTTTTTC
GGTTATCAGTCCGAAGAGTTAAGGTGGAGGTATGCTTCGAAACAGAATCCTATAAAAAAG
300
TGTTCCAACCTTGTATATTATATTCAAAACAGATCAATAAACACATCGTCAGAGTCACAA
ACAAGGTTGGAACATATAATATAAGTTTTGTCTAGTTATTTGTGTAGCAGTCTCAGTGTT
360
TTAGTAATATTCTTGGCAAGAATTGTGCAGCTTTTGGCACCGAGGAATGTTTTCAGGCAC
AATCATTATAAGAACCGTTCTTAACACGTCGAAAACCGTGGCTCCTTACAAAAGTCCGTG
420
TTTCTATTAAAAGTGCGATGAGGAAACTGATGACTCAAGAAATATTTCAGATGAAGACAC
AAAGATAATTTTCACGCTACTCCTTTGACTACTGAGTTCTTTATAAAGTCTACTTCTGTG
480
GTAAAGACGCAAGATTCTTACATCTCCAACTGGACGCAGTCCTTCACCAGATTCTTAATG
CATTTCTGCGTTCTAAGAATGTAGAGGTTGACCTGCGTCAGGAAGTGGTCTAAGAATTAC
540
CTCTGGTGGGTCCCTGACTTGTGTCTCTTGGGGTAAAGACTAATGCTTTTGAACACTGTC
GAGACCACCCAGGGACTGAACACAGAGAACCCCATTTCTGATTACGAAAACTTGTGACAG
600
CTCATGTCTTTCGTCGTTGAAAAAACGTTCAACCCCCAGAGAAGCTGACTATACTGGTGA
GAGTACAGAAAGCAGCAACTTTTTTGCAAGTTGGGGGTCTCTTCGACTGATATGACCACT
660
TCAAACCCAAGACCTTCCGCCTGTGGAACAGGCGCGTTGGCCACCTCAATAACAAATCTT
AGTTTGGGTTGTGGAAGGCGGACACCTTGTCCGCGCAACCGGTGGAGTTATTGTTTAGAA
720
GTGCCACGTATTGCCATTTAACATCCAGTAGTATAAATAGGGTGATTACAGAACTGGGAA
CACGGTGCATAACGGTAAATTGTAGGTCATCATATTTATCCCACTAATGTCTTGACCCTT
780

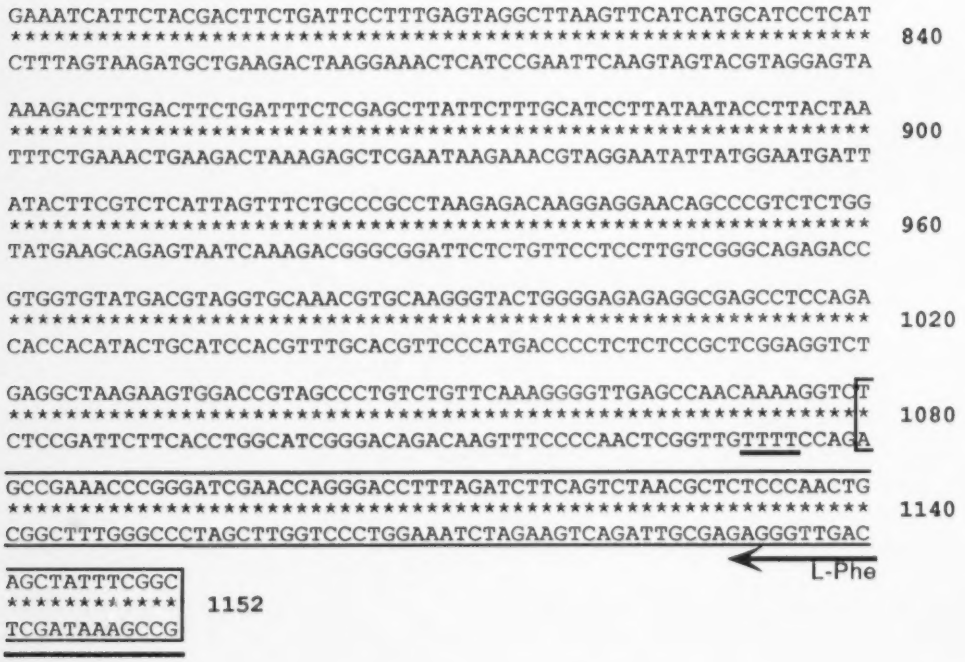
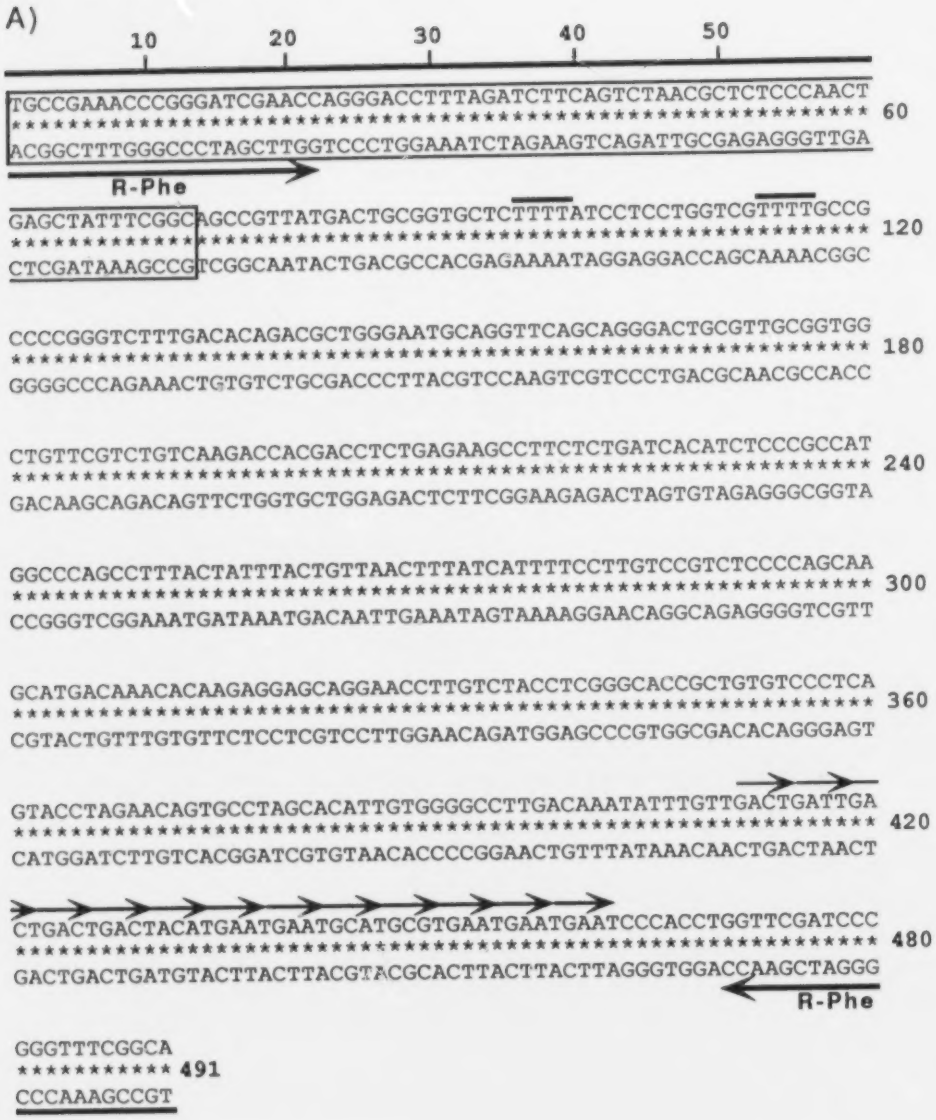
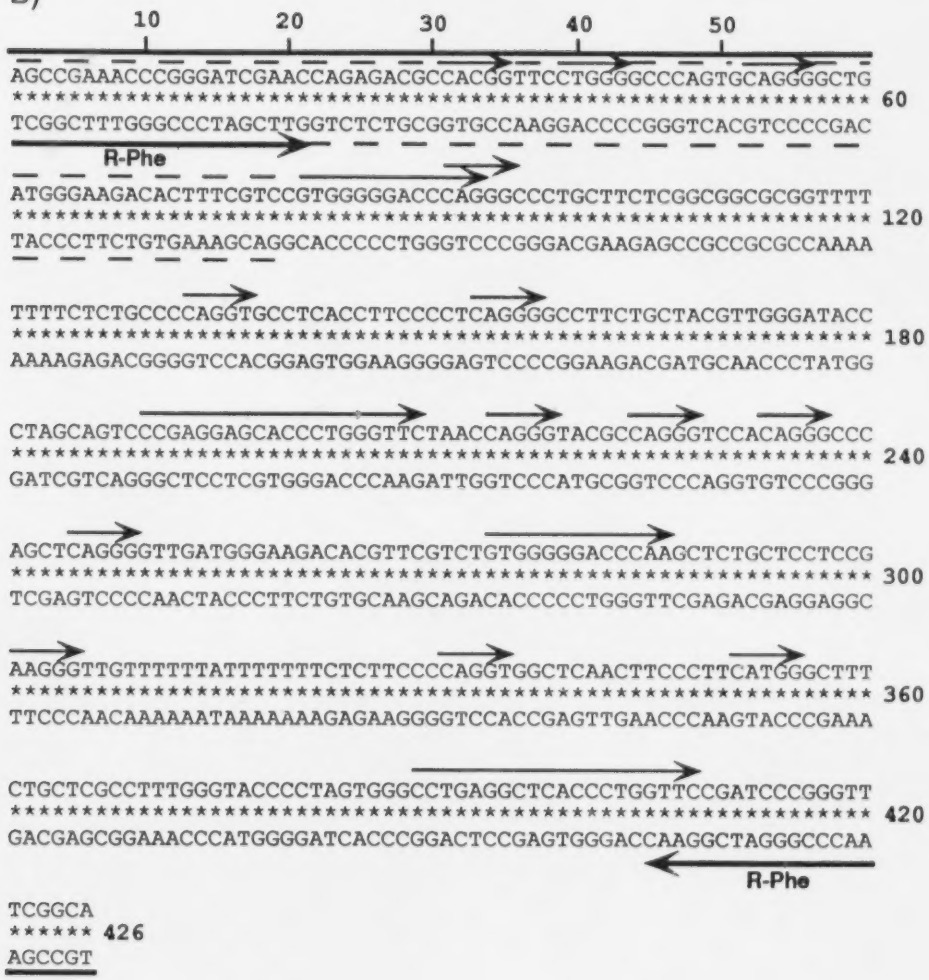

GAAATCATTCTACGACTTCTGATTCCTTTGAGTAGGCTTAAGTTCATCATGCATCCTCAT
840
CTTTAGTAAGATGCTGAAGACTAAGGAAACTCATCCGAATTCAAGTAGTACGTAGGAGTA
AAAGACTTTGACTTCTGATTTCTCGAGCTTATTCTTTGCATCCTTATAATACCTTACTAA
900
TTTCTGAAACTGAAGACTAAAGAGCTCGAATAAGAAACGTAGGAATATTATGGAATGATT
ATACTTCGTCTCATTAGTTTCTGCCCGCCTAAGAGACAAGGAGGAACAGCCCGTCTCTGG
960
TATGAAGCAGAGTAATCAAAGACGGGCGGATTCTCTGTTCCTCCTTGTCGGGCAGAGACC
GTGGTGTATGACGTAGGTGCAAACGTGCAAGGGTACTGGGGAGAGAGGCGAGCCTCCAGA
1020
CACCACATACTGCATCCACGTTTGCACGTTCCCATGACCCCTCTCTCCGCTCGGAGGTCT
GAGGCTAAGAAGTGGACCGTAGCCCTGTCTGTTCAAAGGGGTTGAGCCAACAAAAGGTCT
1080
CTCCGATTCTTCACCTGGCATCGGGACAGACAAGTTTCCCCAACTCGGTTGTTTTCCAGA
GCCGAAACCCGGGATCGAACCAGGGACCTTTAGATCTTCAGTCTAACGCTCTCCCAACTG
1140
CGGCTTTGGGCCCTAGCTTGGTCCCTGGAAATCTAGAAGTCAGATTGCGAGAGGGTTGAC
L-Phe
AGCTATTTCGGC
1152
TCGATAAAGCCG

Figure 32 - Nucleotide sequence of two human DNA fragments contained within pHf7-31 and pHf7-30.

The human DNA fragments found within these recombinants were obtained from tRNA PCR using only the $tRNA^{Phe}$ 3' gene-specific primer. One terminus of the 491 bp insert in pHf7-31 (diagram A) contains a complete human $tRNA^{Phe}$ gene while the other terminus consists of sequence derived only from the 3' gene-specific primer region and is devoid of other tRNA gene sequence. The lower strand of the boxed region encompassing the tRNA gene is the non-coding (tRNA-like) sequence. The plasmid pHf7-30 (diagram B) contains a 426 bp insert that has only primer binding sequences at each terminus and does not contain a complete human $tRNA^{Phe}$ gene. A potential pseudogene is shown surrounded by dashed lines. The primer binding sites (RPhe) are shown as thick arrows. Thin arrows indicate small repeated sequences.


A)
10
20
30
40
50
TGCCGAAACCCGGGATCGAACCAGGGACCTTTAGATCTTCAGTCTAACGCTCTCCCAACT
ACGGCTTTGGGCCCTAGCTTGGTCCCTGGAAATCTAGAAGTCAGATTGCGAGAGGGTTGA
60
R-Phe
GAGCTATTTCGGCAGCCGTTATGACTGCGGTGCTCTTTTATCCTCCTGGTCGTTTTGCCG
CTCGATAAAGCCGTCGGCAATACTGACGCCACGAGAAAATAGGAGGACCAGCAAAACGGC
120
CCCCGGGTCTTTGACACAGACGCTGGGAATGCAGGTTCAGCAGGGACTGCGTTGCGGTGG
GGGGCCCAGAAACTGTGTCTGCGACCCTTACGTCCAAGTCGTCCCTGACGCAACGCCACC
180
CTGTTCGTCTGTCAAGACCACGACCTCTGAGAAGCCTTCTCTGATCACATCTCCCGCCAT
GACAAGCAGACAGTTCTGGTGCTGGAGACTCTTCGGAAGAGACTAGTGTAGAGGGCGGTA
240
GGCCCAGCCTTTACTATTTACTGTTAACTTTATCATTTTCCTTGTCCGTCTCCCCAGCAA
CCGGGTCGGAAATGATAAATGACAATTGAAATAGTAAAAGGAACAGGCAGAGGGGTCGTT
300
GCATGACAAACACAAGAGGAGCAGGAACCTTGTCTACCTCGGGCACCGCTGTGTCCCTCA
CGTACTGTTTGTGTTCTCCTCGTCCTTGGAACAGATGGAGCCCGTGGCGACACAGGGAGT
360
GTACCTAGAACAGTGCCTAGCACATTGTGGGGCCTTGACAAATATTTGTTGACTGATTGA
CATGGATCTTGTCACGGATCGTGTAACACCCCGGAACTGTTTATAAACAACTGACTAACT
420
CTGACTGACTACATGAATGAATGCATGCGTGAATGAATGAATCCCACCTGGTTCGATCCC
GACTGACTGATGTACTTACTTACGTACGCACTTACTTACTTAGGGTGGACCAAGCTAGGG
480
R-Phe
GGGTTTCGGCA
CCCAAAGCCGT
491

B)


10
20
30
40
50
AGCCGAAACCCGGGATCGAACCAGAGACGCCACGGTTCCTGGGGCCCAGTGCAGGGGCTG
60
TCGGCTTTGGGCCCTAGCTTGGTCTCTGCGGTGCCAAGGACCCCGGGTCACGTCCCCGAC
R-Phe
ATGGGAAGACACTTTCGTCCGTGGGGGACCCAGGGCCCTGCTTCTCGGCGGCGCGGTTTT
120
TACCCTTCTGTGAAAGCAGGCACCCCCTGGGTCCCGGGACGAAGAGCCGCCGCGCCAAAA
TTTTCTCTGCCCCAGGTGCCTCACCTTCCCCTCAGGGGCCTTCTGCTACGTTGGGATACC
180
AAAAGAGACGGGGTCCACGGAGTGGAAGGGGAGTCCCCGGAAGACGATGCAACCCTATGG
CTAGCAGTCCCGAGGAGCACCCTGGGTTCTAACCAGGGTACGCCAGGGTCCACAGGGCCC
240
GATCGTCAGGGCTCCTCGTGGGACCCAAGATTGGTCCCATGCGGTCCCAGGTGTCCCGGG
AGCTCAGGGGTTGATGGGAAGACACGTTCGTCTGTGGGGGACCCAAGCTCTGCTCCTCCG
300
TCGAGTCCCCAACTACCCTTCTGTGCAAGCAGACACCCCCTGGGTTCGAGACGAGGAGGC
AAGGGTTGTTTTTTATTTTTTTCTCTTCCCCAGGTGGCTCAACTTCCCTTCATGGGCTTT
360
TTCCCAACAAAAAATAAAAAAAGAGAAGGGGTCCACCGAGTTGAACCCAAGTACCCGAAA
CTGCTCGCCTTTGGGTACCCCTAGTGGGCCTGAGGCTCACCCTGGTTCCGATCCCGGGTT
420
GACGAGCGGAAACCCATGGGGATCACCCGGACTCCGAGTGGGACCAAGGCTAGGGCCCAA
R-Phe
TCGGCA
426
AGCCGT

C. $tRNA^{Lys}$: The results of PCR using 5' and 3' gene-specific primers for $tRNA^{Lys}$ are shown in Figure 33. Amplification products were screened by hybridizing to the 3' gene-specific primer that was used in the PCR. No detectable amplification products were observed for the single primer PCR involving the 5' gene-specific primer. When the 3' gene-specific primer was used in a PCR, however, at least 6 very faint bands of ~1.2 kb to ~4.5 kb in length were detected by ethidium bromide staining. Two of these bands were weakly detected when probed with the 3' gene-specific primer. Probably a longer exposure time would have allowed visualization of the other bands.

D. $tRNA^{Arg}$: PCR was performed using primers specific for the 5' and 3' conserved regions for one particular $tRNA^{Arg}$ gene sequence. Amplification products were size fractionated on a 2% agarose gel and the DNA subsequently transferred to nylon membrane. This was probed separately with each of the primers used in the PCRs. The results are shown in Figure 34. Amplification from the 5' gene-specific primer alone resulted in the production of a single strong band of ~1.7 kb in length. Two smaller fragments of ~0.25 and 0.35 kb in length were also detected by hybridization. These fragments were detected by hybridizing with the 5' gene-specific primer but not with the 3' gene-specific primer, thus these fragments were probably the result of amplification between two $tRNA^{Arg}$ pseudogenes that appear to lack the 3' region of the tRNA gene sequence. At least five different fragments were amplified in PCR using the 3' gene-specific primer alone. These fragments were ~0.25, 0.40, 0.55, 0.80, and 1.3 kb in length. Three of these fragments (0.25, 0.40, 0.55 kb) hybridized weakly to the 3' gene-specific primer and to the 5' gene-specific primer. It is not clear whether this hybridization is because these fragments contain at least one complete tRNA gene sequence or is due to the result of weak non-specific hybridization. There was no detectable hybridization to the marker DNA that was also present at a much higher concentration on the membrane, however.

Figure 33 - tRNA PCR using primers specific for the human $tRNA^{Lys}$ gene family.

(*Left*) An ethidium bromide stained 1.5% agarose gel showing the amplification products for single primer PCR using only the 5' gene-specific primer (lane A), or the 3' gene-specific primer (lane B). Lane C shows the results of amplification for PCR using both primers. (*Right*) The amplification products were transferred to nylon membrane and hybridized with the $tRNA^{Lys}$ 3' gene-specific primer.

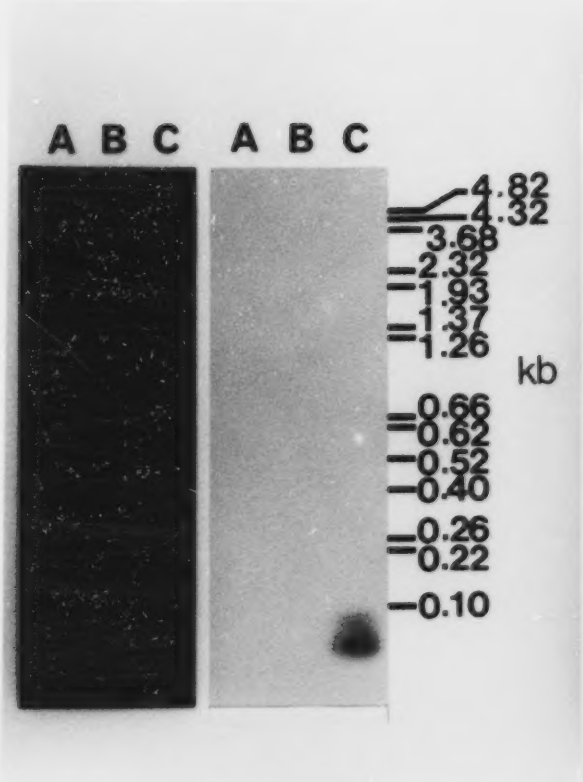
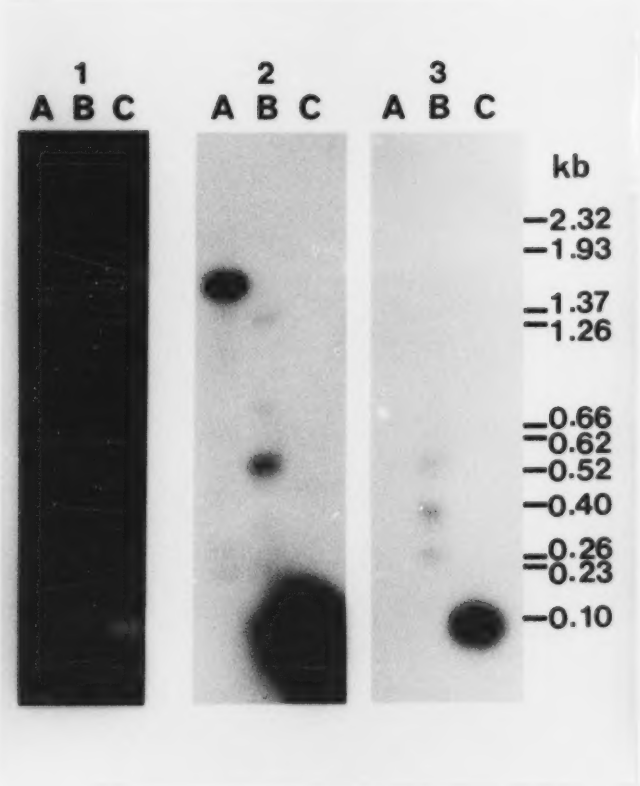

A B C
A B C
4.82
4.32
3.68
2.32
1.93
1.37
1.26
kb
0.66
0.62
0.52
0.40
0.26
0.22
0.10

Figure 34 - tRNA PCR using primers specific for one member of the human $tRNA^{Arg}$ gene family.

Amplification products for single primer PCR using only the 5' gene-specific primer (LArg: lane A), the 3' gene-specific primer (RArg: lane B), or both primers (lane C) were separated by electrophoresis in a 1.5% agarose gel and stained with ethidium bromide (Panel 1). DNA was transferred to nylon membrane and hybridized with the $tRNA^{Arg}$ 5' gene-specific primer (Panel 2) or the $tRNA^{Arg}$ 3' gene-specific primer (Panel 3).


1
A B C
2
A B C
3
A B C
kb
2.32
1.93
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

The ~0.80 and 1.3 kb fragments hybridized only to the 5' gene-specific probe. The only fragment detected for the PCR using both primers was one of approximately 100 bp in length that probably corresponds to the tRNA gene coding sequence alone. When both primers were included in a PCR, none of the other amplification products detected for the single primer PCRs were found.

E. $tRNA^{Trp}$: PCR was performed using primers specific for the 5' and 3' conserved regions of human $tRNA^{Trp}$ genes. The results for PCR using each primer individually or using both primers in a PCR are shown in Figure 35. Amplification using the 5' gene-specific primer alone resulted in the production of at least nine bands. The sizes of these fragments were approximately 0.20, 0.35, 0.40, 0.60, 0.80, 0.90, 1.1, 1.5, and 2.5 kb. Each of these DNA segments were shown to hybridize to the 5' gene-specific primer but not to the 3' gene-specific primer. Conversely, there were no detectable amplification products resulting from PCR using only the 3' gene-specific primer. When both primers were used in a PCR a smear of DNA beginning around 0.10 kb in length was visualized.

F. $tRNA_m^{Met}$: The results of PCR using primers specific for the human $tRNA_m^{Met}$ gene coding sequence are shown in Figure 36. Amplification products were size fractionated on a 2% agarose gel and subsequently transferred to nylon membrane. There were no detectable PCR products from amplification using the human $tRNA_m^{Met}$ 5' gene-specific primer alone in a PCR, as determined by ethidium bromide staining of the gel. While no signal was seen for hybridization of the 3' gene-specific primer, hybridization of the 5' gene-specific primer to the immobilized DNA was observed for a ~0.85 kb fragment and a number of fragments between ~0.2 and 0.5 kb in length. Six fragments of approximately 0.40, 0.60, 0.75, 1.0, 1.8 and 2.2 kb in length were produced from PCR using the human $tRNA^{Met}$ 3' gene-specific primer alone. These fragments, however, did not hybridize to the 5' gene-specific primer. When both of the primers were used in a PCR, a single diffuse band of approximately 0.1 kb was

Figure 35 - tRNA PCR using primers specific for the human $tRNA^{Trp}$ gene family.

Amplification products for single primer PCR using only the 5' gene-specific primer (LTrp: lane A), the 3' gene-specific primer (RTrp: lane B), or both primers (lane C) were separated by electrophoresis in a 1.5% agarose gel and stained with ethidium bromide (Panel 1). DNA was transferred to nylon membrane and hybridized with the $tRNA^{Trp}$ 5' gene-specific primer (Panel 2) or the $tRNA^{Trp}$ 3' gene-specific primer (Panel 3).

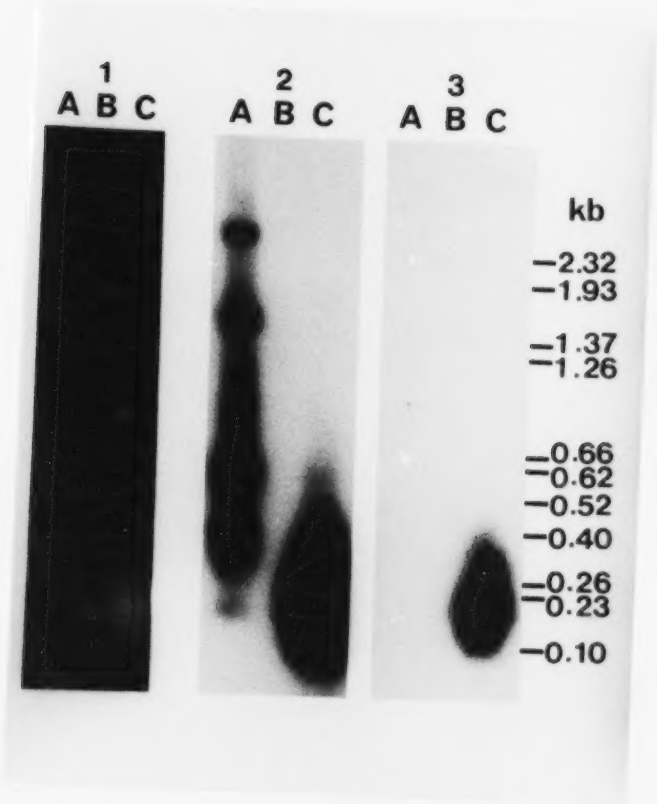
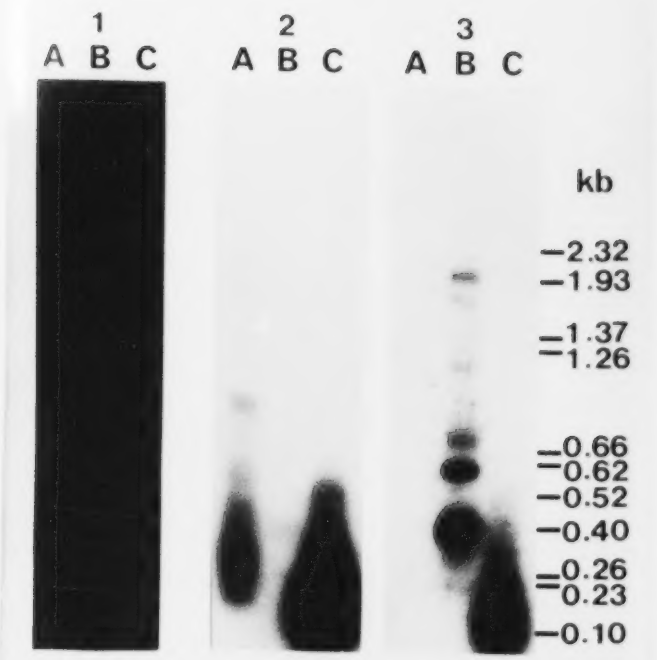

1
A B C
2
A B C
3
A B C
kb
2.32
1.93
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

Figure 36 - tRNA PCR using primers specific for the human $tRNA^{Met}$ gene family.

Amplification products for single primer PCR using only the 5' gene-specific primer (LMet: lane A), the 3' gene-specific primer (RMet: lane B), or both primers (lane C) were separated by electrophoresis in a 1.5% agarose gel and stained with ethidium bromide (Panel 1). DNA was transferred to nylon membrane and hybridized with the $tRNA^{Met}$ 5' gene-specific primer (Panel 2) or the $tRNA^{Met}$ 3' gene-specific primer (Panel 3).


1
A B C
2
A B C
3
A B C
kb
2.32
1.93
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

produced corresponding to amplification between sequences found strictly in the coding sequence for the tRNA gene.

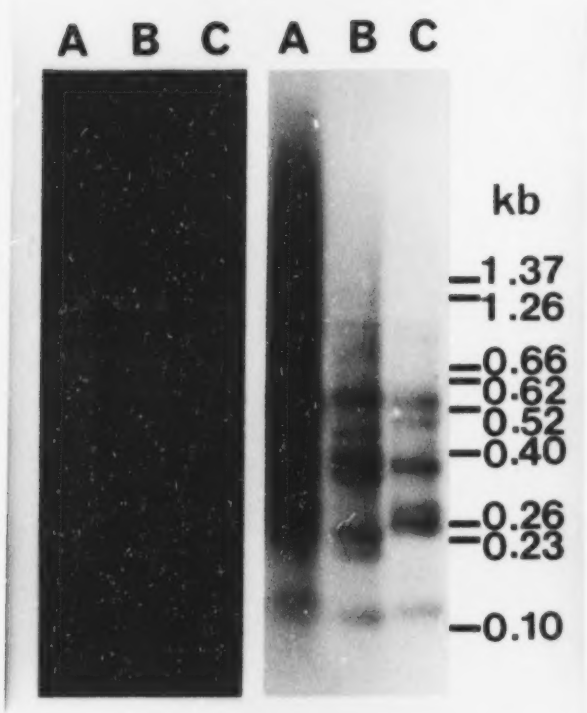
3.3.2 Inverted PCR

A. tRNA^{Phe}: Human genomic DNA was digested with *Cfo*I, *Hin*PI, or *Hae*II. Each restriction digest was diluted serially 10-fold from 10^0 to 10^{-5}, then aliquots of each dilution were used for ligation/circularization of the digested DNA. Inverted PCR was then performed on each dilution using primers that were designed to be specific for the coding sequence of human tRNA^{Phe} genes. Each primer was oriented away from the other primer such that amplification should only be possible if the template DNA had been cut by the restriction enzyme in the flanking regions of the tRNA gene (the two cuts occurring within an amplifiable distance - approx. 4 kb) and only if subsequent circularization of the digested DNA had occurred in the ligation step, thus directing the primers in overlapping directions. Alternatively, amplification might occur if there were repeated tRNA^{Phe} genes occurring within an amplifiable distance. The results of amplification using the tRNA^{Phe} specific primers are shown in Figure 37.

Similar amplification products were observed for the three samples that were initially digested using the different restriction enzymes. All three restriction enzymes used in this case recognize the core sequence 'GCGC', thus it was possible that similar amplification products would be seen. The enzyme *Hae*II recognizes the sequence 'PuGCGCPy', thus it was further expected that the amplification products from this reaction would be a subset of the amplification products seen for those enzymes with four base recognition sequences (*Cfo*I, *Hin*PI). Unexpectedly, the best amplification was seen for the DNA that had not been diluted after digestion (the 10^0 sample). Subsequently, an aliquot of the DNA from the restriction digestion was fractionated on a 0.75% agarose gel to check for the efficiency of digestion. From this it was apparent

Figure 37 - Inverted PCR using primers specific for the human tRNA^{Phe} gene family.

Human genomic DNA was digested with a restriction enzyme and then used directly in a ligation reaction. A sample of the ligation product was then used for inverted PCR. (*Left*) Amplification products from inverse PCR were separated electrophoretically on a 2% agarose gel. The results of PCR of samples where the template was digested with *Hin*PI$_{[G\downarrow CGC]}$ (lane A), *Hae*II$_{[PuGCGC\downarrow Py]}$ (lane B), or *Cfo*I$_{[GCG\downarrow C]}$ (lane C) are shown. (*Right*) DNA was transferred to nylon membrane and hybridized to an oligonucleotide (IR-Phe) which is specific for the 3'-terminus of the human tRNA^{Phe} gene coding sequence that is not found in either primer used in the PCR.


A B C A B C
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

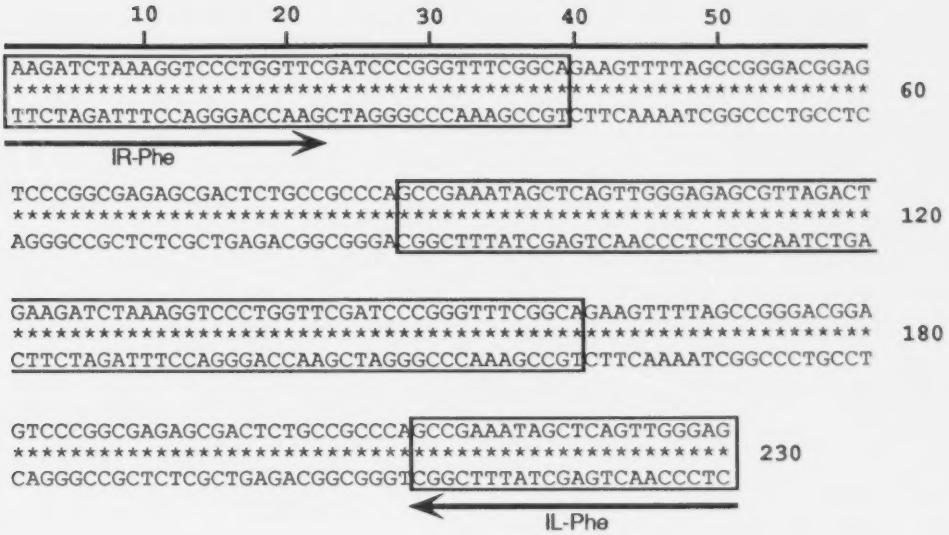
that the DNA was not efficiently digested, as most of the genomic DNA was still greater than 20 kb (data not shown). Ten bands having approximate lengths of 0.11, 0.26, 0.28, 0.38, 0.50, 0.53, 0.62, 0.66, 1.1 and 1.3 kb were produced from amplification of the sample where the ligated template DNA had been initially digested with *Cfo*I. The amplification products from PCR using the *Hin*PI digested DNA were the same as that for *Cfo*I sample except for the presence of an additional fragment of around 0.90 kb in length. The fragments for the *Hae*II digested DNA sample were similar except that the 0.26 and 0.28 kb fragments were absent and two additional bands of 0.21 and 0.23 kb in length were produced. The DNA from these PCRs was transferred to nylon membrane and hybridized to an oligonucleotide that was specific for $tRNA^{Phe}$ sequences that were not found in either of the primer sequences. From this it was evident that some of the amplification products did not contain complete $tRNA^{Phe}$ coding sequences as not all of the fragments were able to hybridize to this probe. In fact, the most predominant signals were due to hybridization of the probe to DNA contained in the fainter fragments visualized by ethidium bromide staining of the amplification products. Hybridization was seen to the ~0.11, 0.26, 0.28, 0.38, 0.50, and 0.53 kb fragments for the amplification products from both *Cfo*I or *HinPI* treated template samples. Additional background hybridization was observed to amplification products from the *Hin*PI treated samples. Clear hybridization was observed for the ~0.10, 0.21, 0.23, 0.37, 0.40, 0.45, and 0.58 kb fragments, and possibly some other larger fragments from the amplification using the *Hae*II treated template DNA.

DNA from each PCR was cloned into the *Hin*cII site of pUC118. After the initial transformation step, positive transformants were screened by colony hybridization. The same oligonucleotide that was used for hybridization to the initial amplification products was used as a probe. A total of 24 positively hybridizing clones were then sequenced. Of these, 11 different sequences were observed. The results of sequencing are shown in Figures 38 to 40.

Figure 38 - Nucleotide sequences of pCfo1, pCfo2, and pCfo35.

Inverted PCR, using $tRNA^{Phe}$ gene-specific primers, was performed on human genomic DNA that was initially digested with *Cfo*I prior to ligation. Amplification products were cloned into the *Hin*cII site of pUC118 for sequence analysis. The complete nucleotide sequence of pCfo1 is shown in (A). A sequence comparison of pCfo2 and pCfo35 is shown in (B). The primer binding site for IL-Phe in these fragments is truncated. Thick arrows indicate the primer binding sites. Boxed regions are complete or partial $tRNA^{Phe}$ gene sequences.

A)


10
20
30
40
50
AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGAAGTTTTAGCCGGGACGGAG
TTCTAGATTTCCAGGGACCAAGCTAGGGCCCAAAGCCGTCTTCAAAATCGGCCCTGCCTC
60
IR-Phe
TCCCGGCGAGAGCGACTCTGCCGCCCAGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACT
AGGGCCGCTCTCGCTGAGACGGCGGGACGGCTTTATCGAGTCAACCCTCTCGCAATCTGA
120
GAAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGAAGTTTTAGCCGGGACGGA
CTTCTAGATTTCCAGGGACCAAGCTAGGGCCCAAAGCCGTCTTCAAAATCGGCCCTGCCT
180
GTCCCGGCGAGAGCGACTCTGCCGCCCAGCCGAAATAGCTCAGTTGGGAG
CAGGGCCGCTCTCGCTGAGACGGCGGGTCGGCTTTATCGAGTCAACCCTC
230
IL-Phe

B)

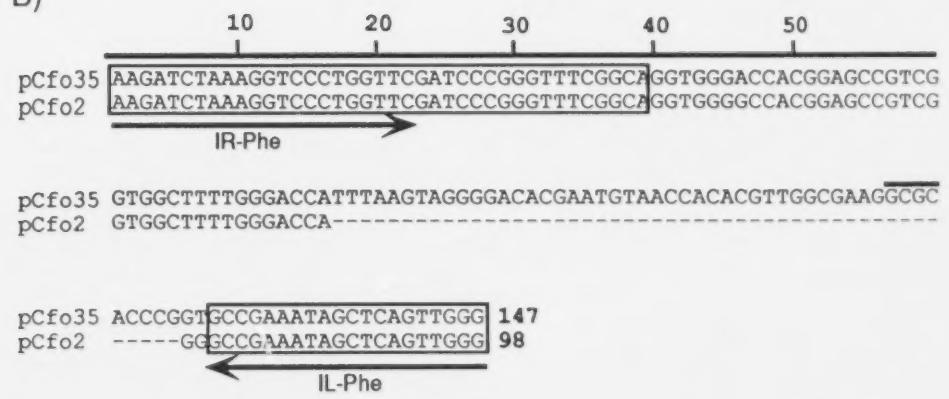

10
20
30
40
50
pCfo35 AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGGTGGGACCACGGAGCCGTCG
pCfo2 AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGGTGGGGCCACGGAGCCGTCG
IR-Phe
pCfo35 GTGGCTTTTGGGACCATTTAAGTAGGGGACACGAATGTAACCACACGTTGGCGAAGGCGC
pCfo2 GTGGCTTTTGGGACCA----------
pCfo35 ACCCGGTGCCGAAATAGCTCAGTTGGG 147
pCfo2 -----GGGCCGAAATAGCTCAGTTGGG 98
IL-Phe

Figure 39 - Nucleotide sequence comparisons of the tRNA^{Phe} gene-containing recombinants pHae2, pHae19, pHae20 and pHae23.

Inverted PCR, using tRNA^{Phe} gene-specific primers, was performed on human genomic DNA which was initially digested with *Hae*II prior to ligation. Amplification products were cloned into the *Hin*cII site of pUC118 for sequence analysis. Complete sequences are shown. Thick arrows indicate primer binding sequences. Multiple primer binding sites are shown for the primer IL-Phe. Not all the IR-Phe binding sites are indicated. Boxed regions represent complete or partial tRNA^{Phe} gene sequences.

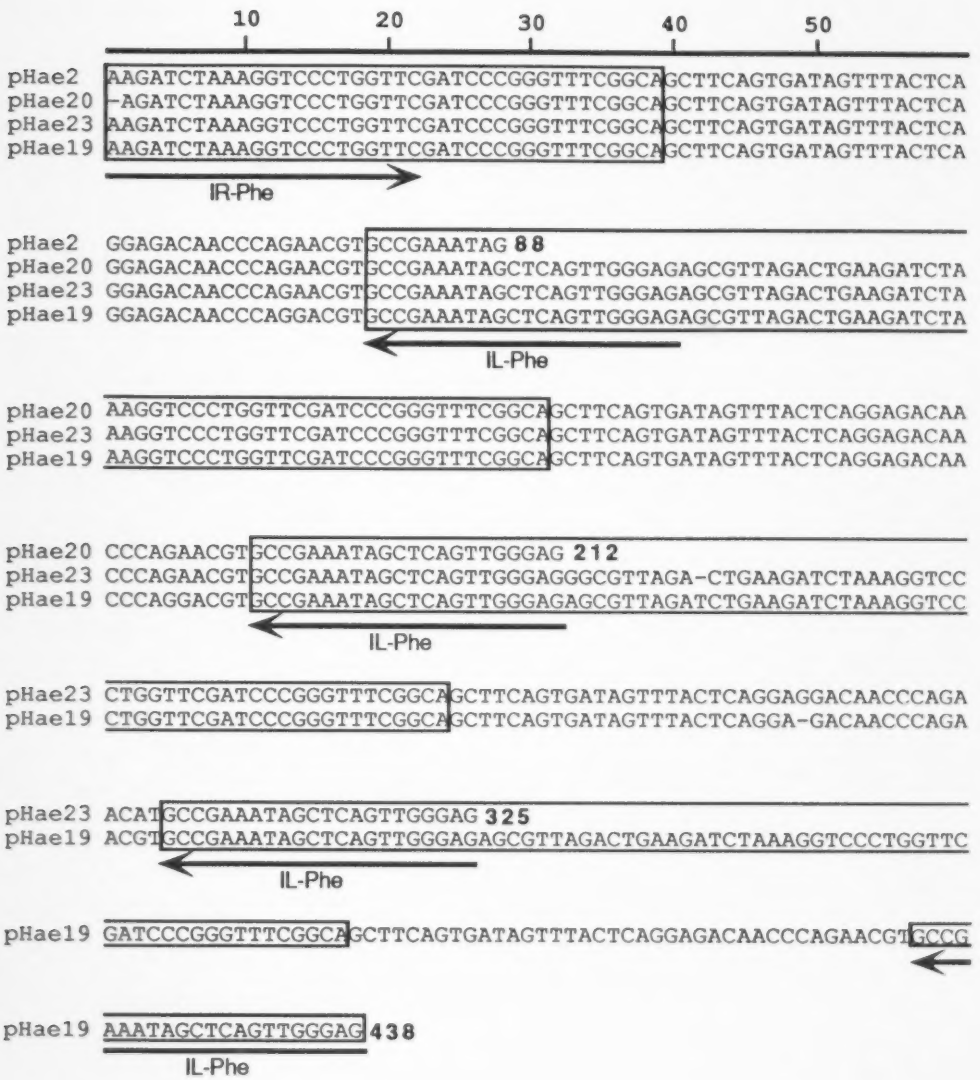
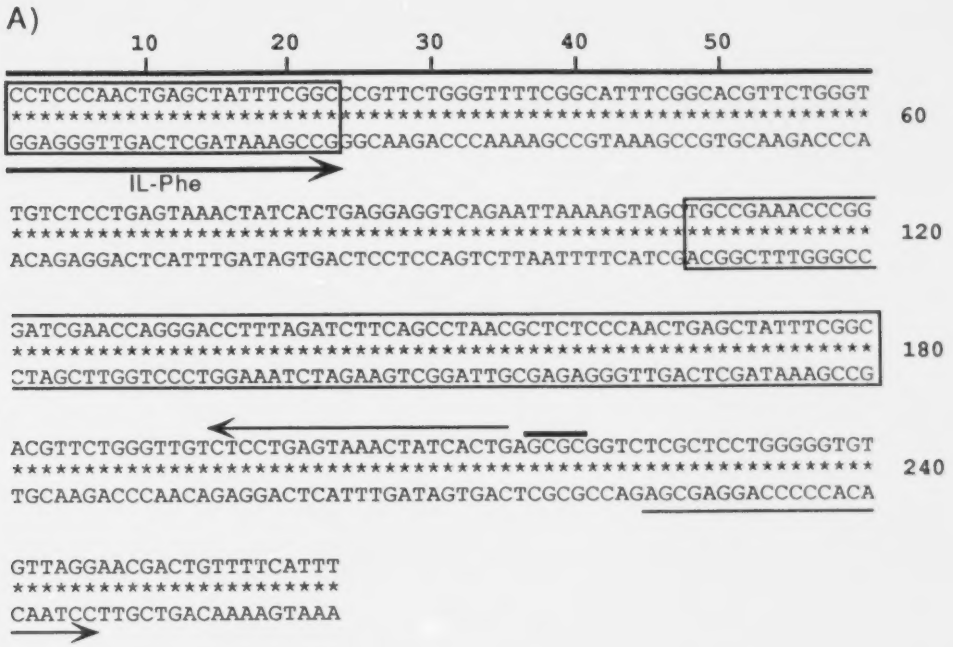
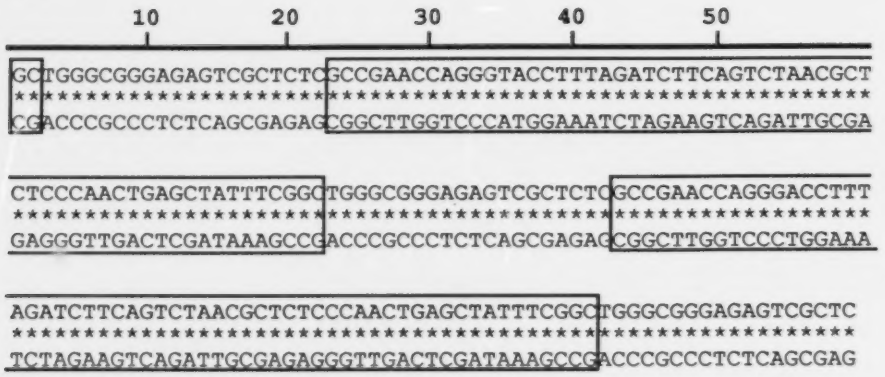

10
20
30
40
50
pHae2 AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCA
pHae20 -AGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCA
pHae23 AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCA
pHae19 AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCA
IR-Phe
pHae2 GGAGACAACCCAGAACGTGCCGAAATAG 88
pHae20 GGAGACAACCCAGAACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTA
pHae23 GGAGACAACCCAGAACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTA
pHae19 GGAGACAACCCAGGACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTA
IL-Phe
pHae20 AAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGAGACAA
pHae23 AAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGAGACAA
pHae19 AAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGAGACAA
pHae20 CCCAGAACGTGCCGAAATAGCTCAGTTGGGAG 212
pHae23 CCCAGAACGTGCCGAAATAGCTCAGTTGGGAGGGCGTTAGA-CTGAAGATCTAAAGGTCC
pHae19 CCCAGGACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGATCTGAAGATCTAAAGGTCC
IL-Phe
pHae23 CTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGAGGACAACCCAGA
pHae19 CTGGTTCGATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGA-GACAACCCAGA
pHae23 ACATGCCGAAATAGCTCAGTTGGGAG 325
pHae19 ACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTAAAGGTCCCTGGTTC
IL-Phe
pHae19 GATCCCGGGTTTCGGCAGCTTCAGTGATAGTTTACTCAGGAGACAACCCAGAACGTGCCG
pHae19 AAATAGCTCAGTTGGGAG 438
IL-Phe

Figure 40 - Nucleotide sequence of the $tRNA^{Phe}$ genes or pseudogenes within pHnp2, pHnp4, pHnp11, and pHnp12.

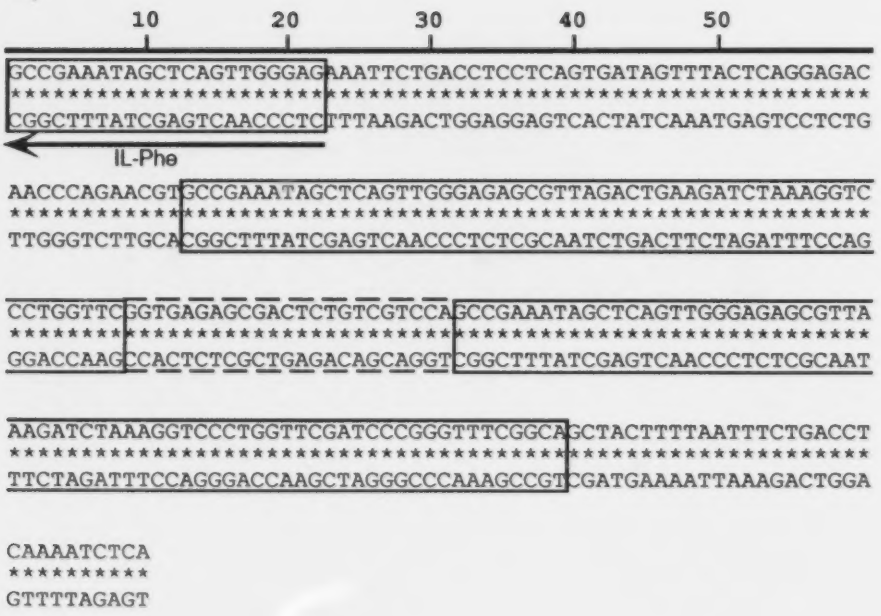
Inverted PCR, using $tRNA^{Phe}$ gene-specific primers, was performed on human genomic DNA which was initially digested with *Hin*PI prior to ligation. Amplification products were cloned into the *Hinc*II site of pUC118 for sequence analysis. Partial sequences of pHnp4 (A), pHnp2 (C), and pHnp11 (D) are shown. The complete sequence of pHnp12 is shown in (B). Boxed regions indicate partial or complete $tRNA^{Phe}$ genes or pseudogenes. The boxed regions in pHnp2 are probably pseudogenes. A probable primer dimer artifact is shown in the sequence of pHnp12. The primer binding site within pHnp11 is inverted relative to what is expected. Also, pHnp11 demonstrates a $tRNA^{Phe}$ gene-like sequence (boxed region plus dashed lines) that ends at the 5' terminus of a complete $tRNA^{Phe}$ gene. Thick arrows represent primer binding sites. Thin arrows indicate primers that were used in subsequent PCRs from human genomic DNA. Solid bars indicate the restriction recognition sequence of the enzyme that was used in the initial digestion step for inverse PCR.


A)
10
20
30
40
50
CCTCCCAACTGAGCTATTTCGGCCCGTTCTGGGTTTTCGGCATTTCGGCACGTTCTGGGT
GGAGGGTTGACTCGATAAAGCCGGGCAAGACCCAAAAGCCGTAAAGCCGTGCAAGACCCA
60
IL-Phe
TGTCTCCTGAGTAAACTATCACTGAGGAGGTCAGAATTAAAAGTAGCTGCCGAAACCCGG
ACAGAGGACTCATTTGATAGTGACTCCTCCAGTCTTAATTTTCATCGACGGCTTTGGGCC
120
GATCGAACCAGGGACCTTTAGATCTTCAGCCTAACGCTCTCCCAACTGAGCTATTTCGGC
CTAGCTTGGTCCCTGGAAATCTAGAAGTCGGATTGCGAGAGGGTTGACTCGATAAAGCCG
180
ACGTTCTGGGTTGTCTCCTGAGTAAACTATCACTGAGCGCGGTCTCGCTCCTGGGGGTGT
TGCAAGACCCAACAGAGGACTCATTTGATAGTGACTCGCGCCAGAGCGAGGACCCCCACA
240
GTTAGGAACGACTGTTTTCATTT
CAATCCTTGCTGACAAAAGTAAA
B)
10
20
30
40
50
AAGATCTAAAGGTCCCTGGAAGATCTAAAGGTCCCTGGTTCAAGCGATTCTCCTGCCTCA
TTCTAGATTTCCAGGGACCTTCTAGATTTCCAGGGACCAAGTTCGCTAAGAGGACGGAGT
60
IR-Phe
GTCTCTTCCAGTAGGTGGGATTACAGGCGCGCTGTGTTGAAGGAGTGACAATTATGCTAA
CAGAGAAGGTCATCCACCCTAATGTCCGCGCGACACAACTTCCTCACTGTTAATACGATT
120
AACCAAAATGCAACTGTCGAAACCCGGGATTGAACCAGGGACCTTTAGATCTTCAGTCTA
TTGGTTTTACGTTGACAGCTTTGGGCCCTAACTTGGTCCCTGGAAATCTAGAAGTCAGAT
180
ACGCTCTCCCAACTGAGCTATTTCGGCACGTTCTGGGTCTGTCTCCTGAGTAAACTATCA
TGCGAGAGGGTTGACTCGATAAAGCCGTGCAAGACCCAGACAGAGGACTCATTTGATAGT
240
CTGAGTGTGCCGAAACCCGGGATCGAACCAGGGACCTCTTAGTATCTT
GACTCACACGGCTTTGGGCCCTAGCTTGGTCCCTGGAGAATCATAGAA
288
IR-Phe

C)


10
20
30
40
50
GCTGGGCGGGAGAGTCGCTCTCGCCGAACCAGGGTACCTTTAGATCTTCAGTCTAACGCT
CGACCCGCCCTCTCAGCGAGAGCGGCTTGGTCCCATGGAAATCTAGAAGTCAGATTGCGA
CTCCCAACTGAGCTATTTCGGCTGGGCGGGAGAGTCGCTCTCGCCGAACCAGGGACCTTT
GAGGGTTGACTCGATAAAGCCGACCCGCCCTCTCAGCGAGAGCGGCTTGGTCCCTGGAAA
AGATCTTCAGTCTAACGCTCTCCCAACTGAGCTATTTCGGCTGGGCGGGAGAGTCGCTC
TCTAGAAGTCAGATTGCGAGAGGGTTGACTCGATAAAGCCGACCCGCCCTCTCAGCGAG

D)

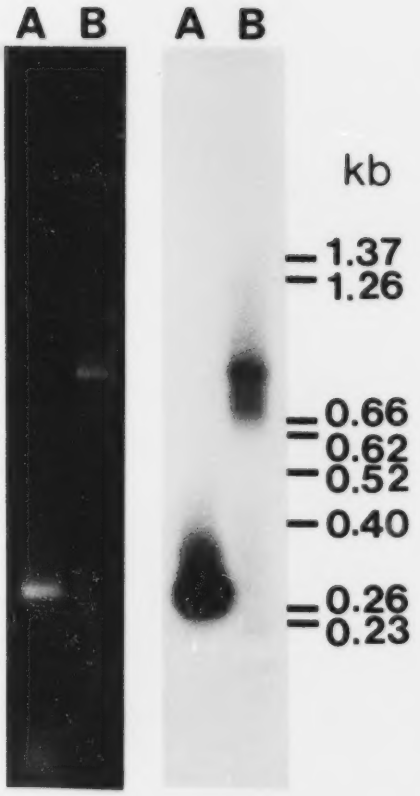

10
20
30
40
50
GCCGAAATAGCTCAGTTGGGAGAAATTCTGACCTCCTCAGTGATAGTTTACTCAGGAGAC
CGGCTTTATCGAGTCAACCCTCTTTAAGACTGGAGGAGTCACTATCAAATGAGTCCTCTG
IL-Phe
AACCCAGAACGTGCCGAAATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTAAAGGTC
TTGGGTCTTGCACGGCTTTATCGAGTCAACCCTCTCGCAATCTGACTTCTAGATTTCCAG
CCTGGTTCGGTGAGAGCGACTCTGTCGTCCAGCCGAAATAGCTCAGTTGGGAGAGCGTTA
GGACCAAGCCACTCTCGCTGAGACAGCAGGTCGGCTTTATCGAGTCAACCCTCTCGCAAT
AAGATCTAAAGGTCCCTGGTTCGATCCCGGGTTTCGGCAGCTACTTTTAATTTCTGACCT
TTCTAGATTTCCAGGGACCAAGCTAGGGCCCAAAGCCGTCGATGAAAATTAAAGACTGGA
CAAAATCTCA
GTTTTAGAGT

Many of the clones analysed were shown to contain multiple copies of $tRNA^{Phe}$ genes arranged as tandem repeats. The basic unit length of tRNA gene plus flanking sequence was around 100 bp. Each fragment was shown to contain the 5' gene-specific primer sequence at one terminus and the 3' gene-specific primer at the other. Each fragment was also shown to contain the additional 3' tRNA gene sequence (~15 bp) that was not found in the 3' gene-specific primer sequence (IR-Phe). Many of the longer sequences containing multiple tRNA genes were concatemers of the basic unit length sequence (~100 x n, where n is the number of tRNA genes contained within the sequence, including both primer sequences at the termini as one tRNA gene sequence). Unexpectedly, not all of the sequences analyzed contained the restriction endonuclease recognition sequence of the enzyme used in the original digest. This suggests that many of the amplification products were derived from unligated DNA. Because of this unexpected result, PCR using the same primers was performed on undigested human genomic DNA. The amplification products thus generated appeared to be the same as the most intensely staining fragments seen from those templates that had been digested and ligated prior to PCR. These amplification products from undigested DNA failed to hybridize to the oligonucleotide that was shown to hybridize to the amplification products derived from PCR using, as the template, human genomic DNA that had been digested and ligated (data not shown).

For those sequences that were shown to contain the restriction endonuclease recognition site of the enzyme which was used in the initial digestion step, primers were designed that were oriented in non-overlapping directions and were flanking the restriction site. PCR on human genomic DNA was then performed using these primers. The results are shown in Figure 41. The primers that were derived from the sequence from pHnp12 produced a single band of ~800 bp in length when used in a PCR directly from the human genome. The primers designed from pHnp4 sequence resulted in the amplification of a ~300 bp fragment.

Figure 41 - Amplification products from PCR from human genomic DNA using primers that are complementary to DNA regions flanking the $tRNA^{Phe}$ genes found in pHnp4 and pHnp12.

Amplification products from inverted PCR for the human $tRNA^{Phe}$ gene family were cloned into the *Hinc*II site of pUC118 for sequence analysis (see Figure 41). Two recombinants, pHnp4 and pHnp12, were shown to contain the restriction endonuclease recognition site for the enzyme which was used in the digestion step for inverted PCR. Primers were designed in non-overlapping directions based on the nucleotide sequence of pHnp4 (L-Hnp4, R-Hnp4) and pHnp12 (L-Hnp12, R-Hnp12). Each primer set should be oriented in overlapping directions relative to the sequence found in the human genome. PCR was performed using human genomic DNA as a template. Amplification products from PCR using the pHnp4 specific primers (lane A) and the pHnp12 specific primers (lane B) were separated on a 1% agarose gel and stained with ethidium bromide (*left*). DNA was transferred to nylon membrane and hybridized to a $tRNA^{Phe}$ gene-specific oligonucleotide, IR-Phe (*right*).


A B
A B
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23

These fragments were subsequently shown to hybridize to an oligonucleotide probe (IR-Phe) which was specific for the $tRNA^{Phe}$ gene coding sequence.

The inverted PCR procedure was repeated for $tRNA^{Phe}$ using human genomic DNA that had been digested with *Hin*dIII. Samples were digested and subsequently diluted serially 10 fold from 10^0 to 10^{-5} prior to ligation. PCR was then performed on each dilution. The results are shown in Figure 42. In this example, the most intensely hybridizing fragments were seen for the sample where the template DNA had been diluted 10^{-5} prior to ligation. Visualization of amplification products by ethidium bromide staining was difficult due to the low amount DNA present, however, there appeared to be at least 3-4 bands that were detectable by hybridization with a probe which contained the complete $tRNA^{Phe}$ coding sequence.

B. $tRNA^{His}$: Approximately 5 μg of human genomic DNA was digested with *Hpa*II or *Hin*dIII. This was diluted serially 10 fold from 10^0 to 10^{-5} and aliquots of each dilution were used in a ligation reaction. PCR was performed using histidine tRNA gene-specific primers that were oriented in non-overlapping directions. Amplification products were size fractionated on a 2% agarose gel and the DNA transferred to nylon membrane. Samples were probed with an oligonucleotide that was specific for histidine tRNA gene sequences not found in either of the PCR primer sequences. The results are shown in Figure 43. As was seen for the PCR for $tRNA^{Phe}$, the greatest amplification was detected when the template DNA was undiluted. Surprisingly, when the amplification products were visualized by ethidium bromide staining there were many similar bands in both the *Hpa*II and *Hin*dIII digested template samples. However, the hybridization results indicate that amplification was not the same in both cases. The *Hpa*II digested samples show 4 bands of between 0.4 and 1.0 kb in length when probed, whereas the *Hin*dIII digested samples showed only 1 strongly hybridizing band of approximately 2 kb in length.

Figure 42 - Inverted PCR using primers specific for the human $tRNA^{Phe}$ gene family.

Approximately 1 µg human genomic DNA was digested with *Hin*dIII and then diluted 10^{-5} fold. This was used directly in a ligation reaction. An aliquot of the ligation product was subsequently used for inverted PCR. Amplification products for inverted PCR were separated on a 2% agarose gel and stained with ethidium bromide (lane A). DNA was transferred to nylon membrane and hybridized to an oligonucleotide (RPhe) that is specific for the 3' terminus of the human $tRNA^{Phe}$ gene-coding sequence (lane B).

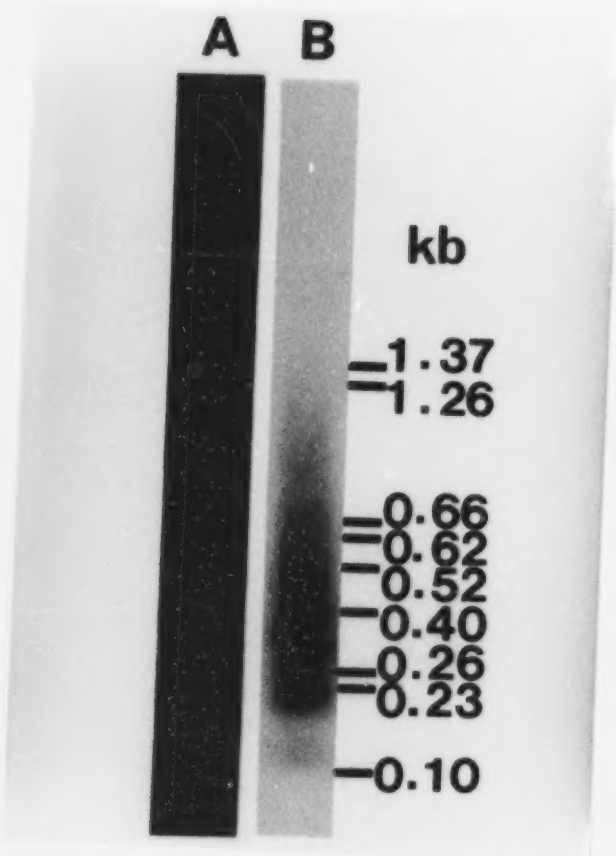
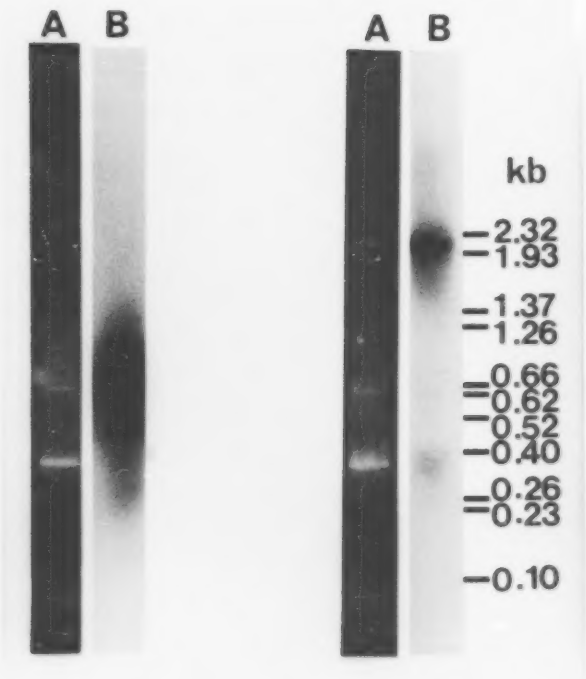

A
B
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

Figure 43 - Inverted PCR using primers specific for the human $tRNA^{His}$ gene family.

Human genomic DNA was digested with *Hpa*II or *Hin*dIII. Digested DNA was 10-fold serially diluted from 10^0 to 10^{-5} and used for ligation reactions. Samples of each ligation reaction were used for inverted PCR. The results for inverted PCR performed using the 10^0 diluted samples for *Hpa*II digested DNA (*left*) and (*Hin*dIII) digested DNA (*right*) are shown. Amplification products were separated on a 2% agarose gel and stained with ethidium bromide (lane A) DNA was subsequently transferred to nylon membrane and hybridized to an oligonucleotide (RHis) that is specific for the 3' region of the $tRNA^{His}$ gene (lane B).


A
B
A
B
kb
2.32
1.93
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

C. tRNA^{Trp}: PCR was done using the conditions described for the histidine tRNA gene family except that the primers were specific for the tryptophan tRNA gene family. The results of this PCR are shown in Figure 44. In both the *Hpa*II and *Hin*dIII digest samples, the strongest amplification occurred for the samples that were not diluted prior to the ligation step. Additionally, in both cases there was a similar pattern of amplification as determined by ethidium bromide staining. In the *Hpa*II digested sample, there were 3 fragments of approximately 0.3 to 0.7 kb in length amplified that were shown to hybridize strongly to the tRNA^{Trp} gene-specific probe. There was no detectable hybridization to amplification products for those samples that had originally been digested with *Hin*dIII.

3.3.3 Alu-tRNA PCR

A. tRNA^{Lys}: PCR was performed using four primers specific for different regions of the human Alu repeat sequence and a primer specific for the conserved sequence of the 5' region of the human tRNA^{Lys} gene sequence. The first stage PCR was performed using a biotinylated tRNA gene-specific primer. The amplification products from this PCR were subjected to streptavidin selection and the PCR repeated using the same primer combinations except that, in this second stage PCR, the gene-specific primer was not biotinylated. Aliquots from the first stage PCR (unselected) and second stage PCR (selected) were fractionated on a 1.5% agarose gel. Additionally, control samples for the PCR using only the Alu specific primers were checked. The amplification products were hybridized to an oligonucleotide that was specific for the 3' region of the tRNA gene. The results of Alu-tRNA PCR for the human tRNA^{Lys} gene family are shown in Figure 45. When only the Alu specific primers were used in a PCR there was a smear of DNA produced. Additionally, when the primer Alu4 was used individually in a PCR, bands of about 220 and 250 bp in length were produced. This was probably due to amplification between these sequences contained within each

Figure 44 - Inverted PCR using primers specific for the human $tRNA^{Trp}$ gene family.

Human genomic DNA was digested with *Hpa*II or *Hin*dIII. Digested DNA was 10-fold serially diluted from 10^0 to 10^{-5} and used for ligation reactions. Samples of each ligation reaction were used for inverted PCR. The results for inverted PCR performed using the 10^0 diluted samples for *Hpa*II digested DNA (*left*) and *Hin*dIII digested DNA (*right*) are shown. Amplification products were separated on a 2% agarose gel and stained with ethidium bromide (lane A) DNA was subsequently transferred to nylon membrane and hybridized to an oligonucleotide (RTrp) which is specific for the 3' region of the $tRNA^{Trp}$ gene (lane B).

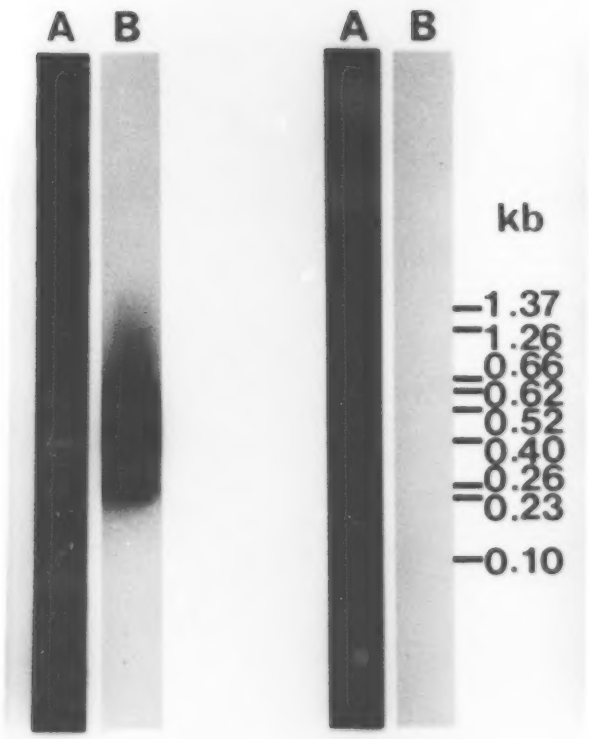
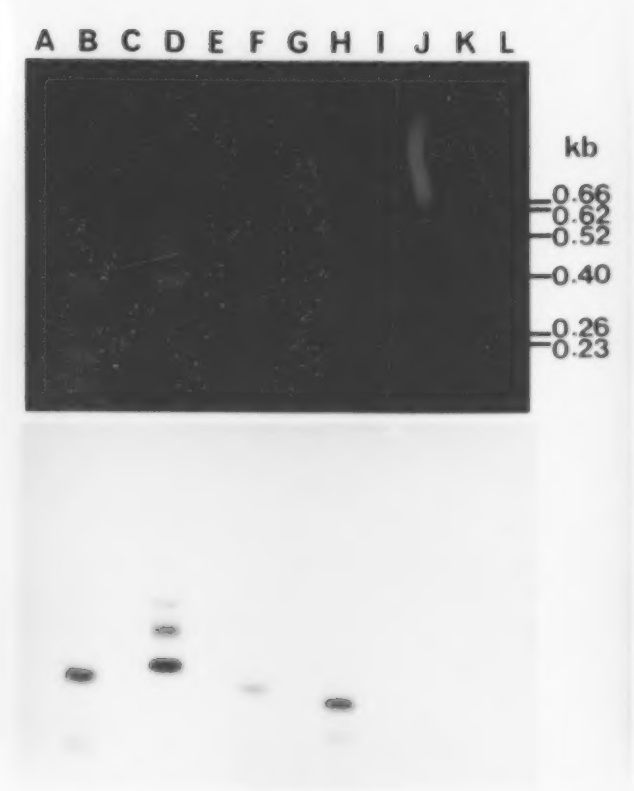

A
B
A
B
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

Figure 45 - Alu-tRNA PCR using primers that are specific for the human $tRNA^{Lys}$ gene family.

A first stage PCR was performed on human genomic DNA using a biotinylated primer that is specific for the 5' region of the human $tRNA^{Lys}$ gene sequence (b-LLys) and four different primers (Alu1, Alu2, Alu3, and Alu4) that are specific for different conserved regions of the human Alu repeat sequence element (lanes A, C, E, and G, respectively). DNA from these PCRs was selected by incubation with streptavidin-agarose. The selected DNA was then used as template for second stage PCRs using a non-biotinylated $tRNA^{Lys}$ 5' gene-specific primer (LLys) with the same combinations of Alu specific primers (lanes B, D, F, H, respectively). Single primer PCR was also performed on human genomic DNA using only the Alu specific primers; Alu1, Alu2, Alu3 and Alu4 (lanes I, J, K, and L, respectively). Amplification products were separated on a 1.5% agarose gel and then stained with ethidium bromide (*top*). DNA was then transferred to nylon membrane and hybridized to an oligonucleotide that is specific for the 3' region of the $tRNA^{Lys}$ gene sequence (RLys). A schematic representation of the Alu-specific primers is shown in Figure 5. A schematic representation of the tRNA-specific primers is shown in Figure 8.

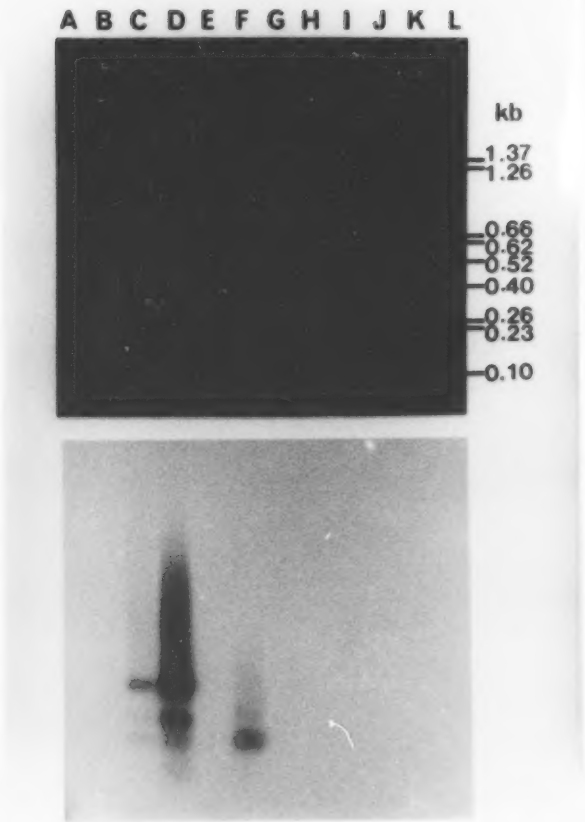

A B C D E F G H I J K L
kb
0.66
0.62
0.52
0.40
0.26
0.23

half of the Alu repeat sequence. There was no detectable hybridization to these control lanes when probed with the $tRNA^{Lys}$ 3' gene-specific primer. PCR performed on the selected DNA samples resulted in very intense amplification products as compared to each of the samples for those PCRs that were performed on unselected genomic DNA. Correspondingly, the hybridization intensity was significantly increased for those PCR samples that had the tRNA gene sequences enriched by selection prior to the PCR. The approximate sizes of the hybridizing fragments for each PCR are: (1) Alu1 + LLys - 0.20 and 0.38 kb, (2) Alu2 + LLys - 0.40, 0.50, and 0.60 kb, (3) Alu3 + LLys - 0.32 kb, and (4) Alu4 + LLys - 0.20 and 0.30 kb.

B. $tRNA^{Phe}$: Alu-tRNA PCR was also performed using a tRNA gene-specific primer complementary to the 5' region of the human $tRNA^{Phe}$ gene family. The results of amplification using the Alu specific primers alone or in conjunction with the $tRNA^{Phe}$ specific primers (biotinylated and non-biotinylated) on non-selected human genomic DNA or amplified DNA that had been selected for $tRNA^{Phe}$ sequences is shown in Figure 46. Again the control amplification reactions using only the Alu specific primers show a smear of DNA that does not hybridize to the 3' gene-specific $tRNA^{Phe}$ primer. Those PCRs that were performed using template DNA that had been enriched for $tRNA^{Phe}$ sequences showed more specific amplification compared to the PCRs that were performed directly from the human genome. The enhancement of specific amplification was not as pronounced as that seen for the $tRNA^{Lys}$ examples as there was still a significant degree of background amplification. There was, however, one primer combination that resulted in a significant amount of specific amplification that was more intense than the background smear due to Alu-Alu amplification. The PCR using the primers Alu2 + LPhe resulted in the production of at least 6 fragments of approximately 0.10, 0.20, 0.30, 0.40, 0.42, 0.50 and 0.55 kb in length that were shown to hybridize to the $tRNA^{Phe}$ 3' gene-specific primer.

Figure 46 - Alu-tRNA PCR using primers that are specific for the human $tRNA^{Phe}$ gene family.

PCR was performed on human genomic DNA using a biotinylated primer that is specific for the 5' region of the human $tRNA^{Phe}$ gene sequence (b-LPhe) and four different primers (Alu1, Alu2, Alu3, and Alu4) that are specific for different regions of the consensus sequence for the human Alu repeat sequence family (lanes A, C, E, and G, respectively). DNA from these PCRs was selected by incubation with streptavidin-agarose. The selected DNA was then used as template for PCRs using a non-biotinylated $tRNA^{Phe}$ 5' gene-specific primer (LPhe) with the same combinations of Alu specific primers (lanes B, D, F, and H, respectively). Single primer PCR was also performed on human genomic DNA using only the Alu specific primers; Alu1, Alu2, Alu3 and Alu4 (lanes I, J, K, and L, respectively). Amplification products were separated on a 1.5% agarose gel and then stained with ethidium bromide (*top*). DNA was transferred to nylon membrane and hybridized to an oligonucleotide that is specific for the 3' region of the $tRNA^{Phe}$ gene sequence (RPhe).

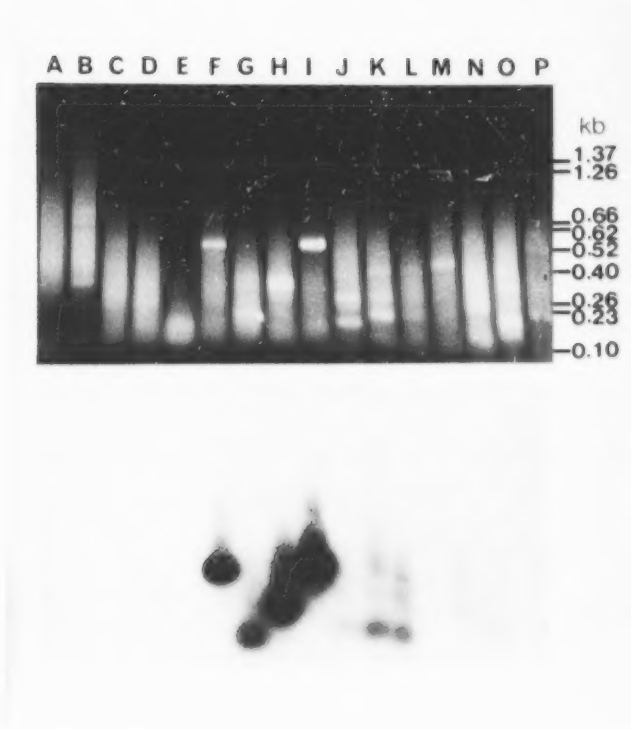

A B C D E F G H I J K L
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

3.3.4 Linker PCR - $tRNA^{Phe}$

In order to amplify novel tRNA gene-containing fragments directly from the human genome, linker PCR was performed where the second primer binding site was supplied by the ligation of linkers to the ends of digested genomic DNA. Human genomic DNA was digested with various restriction enzymes that generate blunt ends (*Eco*RV, *Hpa*I, *Sma*I, *Ssp*I, *Alu*I, *Hae*III, *Rsa*I). Linkers were ligated to the products of these digests. Concatenation of the linker sequences was not possible because their 5' ends were non-phosphorylated. PCR was then performed using one of the linker primers (Plinker1) and either the 5' gene-specific primer or the 3' gene-specific primer. Under these conditions, variable length amplification products were produced (data not shown). The experiment was then repeated using the 5' gene-specific and 3' gene-specific biotinylated primers and the linker primer. These amplification products were selected by mixing with streptavidin agarose. A second round of PCR was then performed using the same primer combinations except that the non-biotinylated tRNA gene-specific primer was used. Additionally, as a control, single primer PCR was performed on the selected DNA that had been previously digested with *Eco*RV or *Hae*III using only the linker primer (Plinker1). Amplification products were fractionated on a 2% agarose gel and the DNA was subsequently transferred to nylon membrane for hybridization to the 3' gene-specific primer for $tRNA^{Phe}$. The results of these PCRs are shown in Figure 47. When visualized by ethidium bromide staining, both reactions using only the linker primer showed a substantial smear of DNA throughout the lanes with a perceptible trace of banding. There was no detectable hybridization of these samples to an oligonucleotide specific for the 3' region of the $tRNA^{Phe}$ gene. In each of the reactions using either of the $tRNA^{Phe}$ specific primers plus the linker primer there was also an obvious smear of DNA present. There was also usually between one and four stronger bands present within the background smear. Good hybridization was observed for amplification products derived from PCR

Figure 47 - Linker PCR for the human tRNA^{Phe} gene family.

Human genomic DNA was digested with a variety of restriction endonucleases that generate blunt ends. Double stranded linkers were ligated to the ends of the digested DNA. A first stage PCR was performed using either a biotinylated 3' tRNA^{Phe} gene-specific primer (b-RPhe) or a 5' tRNA^{Phe} gene specific primer (b-LPhe) with a linker primer (Plinker1). Amplification products were selected for tRNA gene sequences by incubating the DNA with streptavidin-agarose. A second stage PCR was then performed using a linker primer (Plinker1) and either a non-biotinylated tRNA^{Phe} 3' gene-specific primer (RPhe: lanes C to I) or a non-biotinylated tRNA^{Phe} 5' gene-specific primer (LPhe: lanes J to P). The results of amplification using only the linker primer (Plinker1) on human genomic DNA that was initially digested with *Hae*III or *Eco*RV are shown in lanes A and B, respectively. The restriction enzymes used to digest human genomic DNA were *Rsa*I (lanes C and J), *Hae*III (lanes D and K), *Alu*I (lanes E and L), *Ssp*I (lanes F and M), *Sma*I (lanes G and N), *Hpa*I (lanes H and O) and *Eco*RV (lanes I and P). Amplification products were separated on a 1.5% agarose gel and stained with ethidium bromide (*top*). DNA was transferred to nylon membrane and hybridized with a tRNA^{Phe} gene probe prepared by random primer extension.

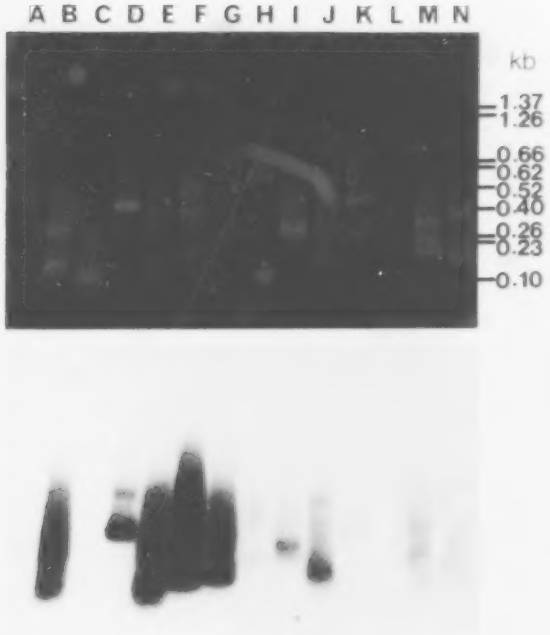

A B C D E F G H I J K L M N O P
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

using the template DNA that was initially digested with *Eco*RV (0.50 kb), *Hpa*I (0.30, 0.32 kb), *Sma*I (0.20 kb) and *Ssp*I (0.50) and were amplified using the 3'-gene-specific primer. There was only very weak hybridization seen to the DNA from the other reaction conditions employed.

In order to increase the total yield and specificity of amplification of tRNA gene sequences and to reduce the amount of background amplification it was decided to attempt the second stage PCR using a 'nested' tRNA gene-specific primer. This nested primer is specific for sequences contained within the $tRNA^{Phe}$ gene that are not contained within the original tRNA gene-specific primer sequence and are found downstream from the first primer. The template used for these reactions was the appropriate selected amplification products from the first stage PCR. Electrophoretic fractionation of amplification products was conducted on 2% agarose gels. The DNA was then transferred to nylon membrane for hybridization analysis (see Figure 48). In all cases there was a noticeable background smear of DNA present, however the yield of the specific bands produced appears to be significantly increased. When hybridized to a probe containing the $tRNA^{Phe}$ gene sequence at 60°C there was strong hybridization to those reactions which utilized the 3' nested primer and only weak hybridization (if any) to the reactions using the 5' nested primer.

Figure 48 - Linker PCR for the human tRNA^{Phe} gene family using nested primers.

Human genomic DNA was digested with a variety of restriction endonucleases that generate blunt ends. Double stranded linkers were ligated to the ends of the digested DNA. A first stage PCR was performed using either a biotinylated 3' tRNA^{Phe} gene-specific primer (b-RPhe) or a 5' tRNA^{Phe} gene-specific primer (b-LPhe). Amplification products were selected for tRNA gene sequences by incubating the DNA with streptavidin-agarose. A second stage PCR was performed on the selected DNA using a linker primer (Plinker1) and either a nested non-biotinylated tRNA^{Phe} 3' gene-specific primer (IR-Phe: lanes A to G) or a nested non-biotinylated tRNA^{Phe} 5' gene-specific primer (IL-Phe: lanes H to N). The tRNA gene-specific primer combinations used were RPhe (first stage) → IL-Phe (second stage) and LPhe (first stage) → IR-Phe (second stage). Schematic representations of the primer sequences used are shown in Figure 8 and Figure 9. The restriction enzymes used for digestion of human genomic DNA were *Eco*RV (lanes A and H), *Hpa*I (lanes B and I), *Sma*I (lanes C and J), *Ssp*I (lanes D and K), *Alu*I (lanes E and L), *Hae*III (lanes F and M) and *Rsa*I (lanes G and N). Amplification products were separated on a 1.5% agarose gel and stained with ethidium bromide (*top*). DNA was transferred to nylon membrane and hybridized with a tRNA^{Phe} gene probe prepared by random primer extension.


A B C D E F G H I J K L M N
kb
1.37
1.26
0.66
0.62
0.52
0.40
0.26
0.23
0.10

4. DISCUSSION

4.1 Standardization of hybridization conditions

One objective of this study was to examine the genomic organization of human tRNA genes. In order to distinguish successfully the chromosomal location of a single human tRNA gene from those of other members of that tRNA gene family, it is necessary to use a specific probe that is substantially different from sequences found at any other location in the human genome. In practice this means first labeling the probe DNA to a very high specific activity to enable detection of that particular sequence from an extremely complex background and second, ensuring that the probe sequence does not contain repeat sequence elements that would hybridize to DNA from other locations within the genome and thus obscure the results of specific hybridization to the desired sequence. Typically, the flanking sequences of human tRNA genes are unique, thus determining the location of a specific tRNA gene locus involves using a probe derived in part from this region. The haploid human genome contains approximately 3×10^9 bp, therefore any hybridization strategy must consider the increased genome complexity of this organism compared to the genomes of other highly studied organisms such as *Escherichia coli* or *Saccharomyces cerevisiae*. This means that for any given mass of genomic DNA used for a hybridization experiment, there are 100 to 1000 fold fewer copies of the target. Thus, probes used for hybridization to human genomic DNA must be labelled to a very high specific activity to allow detection of a hybridization event within a reasonable time of exposure (usually less than 5 days). Moreover, the proportion of any particular human chromosomes found in a given mass of DNA isolated from mouse-human somatic cell hybrids is usually small. Therefore, when dealing with extended exposure times the signal to noise ratio becomes very important. The presence of even a small degree of non-specific hybridization will obscure the results of a weak, specific hybridization. Noise associated with hybridization includes those signals derived from the hybridization of the probe to other similar sequences

within the target DNA and also the non-specific binding of the probe DNA to the solid membrane support.

The conditions necessary for specific hybridization to single copy genes within the human genome were examined. The temperature of hybridization, ionic conditions utilized for hybridization and for washing, probe composition, and the method of probe preparation are the most important factors involved in determining the success of hybridization. Hybridization stringency is relative to the melting temperature (T_m) of a DNA (or RNA) fragment which is determined by the length and base composition of the probe used for hybridization. The conditions of hybridization were investigated for three different probe preparation methods: (1) polynucleotide kinase labeling of oligonucleotide probes, (2) M13 primer extension for M13-human recombinant clones, and (3) random primer probe preparation of DNA fragments excised from low melting point agarose.

Oligonucleotides used as probes are short segments of single stranded DNA, usually 17 to 25 nt in length. One advantage of using an oligonucleotide probe is the ability to precisely define the sequence of the probe, making this method ideal for the detection of target sequences when only limited sequence information is available or when two sequences to be distinguished are very similar precluding the use of longer probe sequences. Because of the limited degree of base pairing between the oligonucleotide probe and the target sequence, the thermodynamic equilibrium of hybridization of these probes is very sensitive to temperature. Using oligonucleotides to determine the chromosomal location of human tRNA genes would be useful for localizing examples where two or more highly homologous but distinct tRNA loci have been described, such as the human $tRNA^{Tyr}$ genes described by van Tol et al., (1987), MacPherson (1988) and the $tRNA^{Tyr}$ genes described in this study. Many of the differences among these sequences occur solely in the intron regions. Also, it was noted that position 16 of some human tyrosine tRNA molecules contains a U rather

than a C as was expected from all known examples of genes coding for this tRNA (Johnson et al., 1985; van Tol et al., 1987). It is possible that this is a template coded polymorphism or it may be the result of a post-transcriptional modification event. In order to discern the difference quickly, it might be possible to utilize differential hybridization to human genomic DNA of two similar oligonucleotides for this region.

The effectiveness of using oligonucleotide probes to detect sequences within the human genome was investigated briefly. The T_m (melting temperature) of a 19 nt long oligonucleotide that was perfectly matched to the sequence contained within a human $tRNA^{Tyr}$ gene was estimated to be ~55° C. While there was no detectable hybridization of the perfectly matched probe to $tRNA^{Tyr}$ sequences contained within the human genome at temperatures of 45°C or 48°C (data not shown), specific hybridization of this probe was observed when the hybridization was performed at 42°C. Approximately 12-14 fragments were observed to hybridize to this probe in *Eco*RI digests of human DNA, with a lesser number of hybridizing fragments observed in *Bam*HI, *Bgl*II and *Hin*dIII digests of human DNA. Similar results were reported by van Tol and Beier, (1988). Thus it would appear that there are at least twelve tyrosine tRNA genes within the human genome and that some of these are contained on the same *Bam*HI, *Bgl*II, or *Hin*dIII fragments. The actual number of $tRNA^{Tyr}$ genes is probably higher than this estimate since it is known that, at least for this tRNA gene family, many of these tRNA genes are tightly clustered (this study; MacPherson and Roy, 1986; MacPherson, 1988). The intensity of hybridization of this probe to human genomic DNA was very weak, even after a 5 day exposure to X-ray film. Thus, under these conditions these probes would be unsuitable for use in examining human chromosomal mapping panels. There was no detectable signal seen for hybridization of the single base mismatched probe to human genomic DNA. Because no positive results were seen for this probe, it is difficult to determine if this means that a C is template

encoded at position 16 of all human $tRNA^{Tyr}$ genes thus causing unstable hybridization.

These results might possibly be improved by the addition of cationic detergents (e.g. cetyl trimethylammonium bromide) to the hybridization solution. These reagents have been shown to increase the rate of homologous pairing of complementary sequences within complex DNA sources (Pontius and Berg, 1991), however suitable conditions for hybridization to membrane immobilized DNA have not yet been reported. However, the rate of base pairing of complementary sequences in my experiments may not be the most important problem. Rather, the relatively low specific activity of the probe may be the critical limiting factor. Similar experiments performed by van Tol and Beier (1988) showed strong specific hybridization of a 24 nt long $tRNA^{Tyr}$ gene-specific oligonucleotide to human genomic DNA. My results might be improved by decreasing the mass of oligonucleotide used for endlabelling, thereby potentially increasing the specific activity of these probes. Also, longer exposure times (>5 days) would have allowed better detection of hybridization signals. Other labeling methods, such as isotope incorporation using terminal transferase, may allow a significant increase in the specific activity of oligonucleotide probes so that they could be used effectively on human genomic DNA sources. Additionally, with the advent of biotinylated phosphoramidites that allow the incorporation of multiple labels into an oligonucleotide sequence, it may be possible to utilize novel non-radioactive methods to increase the effectiveness of detection of single copy genes in human genomic DNA.

In order to increase the specific activity of probes used for detection of tRNA genes, the M13 primer extension method of probe preparation was examined. A single stranded human-M13 recombinant containing a segment of DNA from λHt7 was hybridized to immobilized human DNA at 40° or 60°C (see Figure 10). The probe was prepared by Klenow extension from a primer in the direction of the M13 DNA, leaving the human DNA segment single stranded to allow hybridization to target sequences.

There were noticeable strongly hybridizing fragments seen when this probe was hybridized at 40°C but these were partially obscured by the high background hybridization. This background signal may be due to repeat sequences that are contained within the human DNA segment of the probe, although no Alu repeat sequences were detected within the region sequenced (Doran et al., 1987). Alternatively, it has been reported that M13 bacteriophage DNA contains sequences that may hybridize to repeated sequences within the human genome (Vassart et al., 1987). When the temperature of hybridization was increased to 60°C, the background hybridization was removed but specific hybridization was too weak to allow detection of specific fragments within a 3 day exposure to X-ray film.

The random primer probe preparation method has proven vital to detection of single copy sequences within the human genome. The advantage of this method is that it is possible to label DNA fragments to very high specific activity, however, the reannealing of complementary strands from the double stranded probe DNA is a disadvantage.

The effect of temperature on hybridization of a 610 bp $tRNA^{Gln}$ containing fragment from λHt4 was examined (see Figure 11). This probe was prepared by random primer extension and hybridized at 50°C and 65°C to human genomic DNA. At both temperatures, there were strong signals seen for fragments of ~4.5 and 8.5 kb in length for *Bam*HI and *Bgl*II digests of human genomic DNA, respectively. The effect of temperature on the hybridization of this probe appeared to be minimal, at least for these two temperatures examined. Hybridization to the expected fragments in *Eco*RI and *Hin*dIII digests (0.6 kb for *Eco*RI and two *Hin*dIII fragments of ~0.7 and 2.9 kb) was also seen but these were much weaker. At least eight other hybridizing fragments were detected in these digests. These may represent hybridization of this probe to partially homologous target sequences elsewhere in the human genome or it may be that the temperatures examined do not represent very stringent conditions of

hybridization for probes of this size. Similar observations have been made for $tRNA^{Gly}$ gene-containing fragments (McBride et al., 1989). The extensive cross-hybridization that was observed was explained as the result of homology of the flanking sequences with other regions in the genome, rather than being hybridization strictly to other glycine isoacceptor tRNA genes. These cross-hybridizing regions co-segregated with the $tRNA^{Gly}$ gene loci in chromosomal localization experiments, thus it was suggested that these sequences represent amplification (i.e. tandem duplication) of this DNA segment.

The effect of flanking sequence on the hybridization of probes derived from this same segment of DNA (λHt4: Roy et al., 1982) was analyzed by hybridizing a 72 bp $tRNA^{Gln}$ gene probe (containing no flanking sequence), a 105 bp *Rsa*I-*Hin*dIII fragment (containing 35 bp of 5' flanking sequence from the $tRNA^{Gln}$ gene), and a 610 bp *Eco*RI fragment (containing 540 bp of flanking sequence from the 5' and 3' flanking regions of the $tRNA^{Gln}$ gene) to human genomic DNA digested with *Bam*HI, *Bgl*II, *Eco*RI, or *Hin*dIII (see Figure 12). When the hybridizations were performed at 65°C there was a noticeable difference in the results obtained with the three probes. The shorter probes tended to generate more background hybridization. This may represent hybridization of these probe sequences to many similar sequences present in the genome. It is possible that these other sequences are remnants of the evolution of these tRNA genes or it is possible that human $tRNA^{Gln}$ genes share a considerable amount of sequence similarity with repeated sequence elements contained within the human genome. Certain human tRNA genes have been shown to be very similar to a number of repeated sequence elements and it has also been proposed that for certain human tRNA gene families, the repeat elements actually evolved from these tRNA genes (Daniels and Deininger, 1985; Sakamoto and Okada, 1985). The longer probe sequence generated the most intense hybridization to a few specific bands. This must be due to the sequence that flanks the $tRNA^{Gln}$ gene contained within this probe. The

number and intensity of hybridizing fragments observed for the hybridization of this 610 bp probe under stringent conditions suggests that other homologous sequences are present in the human genome. The intensity of hybridization of the longer probe to the expected fragments was stronger than that seen for the shorter probe sequences. From this it was concluded that for the purpose of hybridization to specific sequences within the human genome (at least for human tRNA gene-containing fragments) a longer probe sequence is more desirable. This desirability of longer probe sequences has to be balanced against the increased probability of containing other repeat sequence elements.

4.2 Human tRNA gene chromosomal mapping studies

It was necessary to determine whether detection of specific hybridization of tRNA gene probes to human sequences could be differentiated from hybridization to similar sequences contained within mouse genomic DNA. To do this, hybridization of tRNA gene-containing probes was done to human and mouse genomic DNA samples digested with *Bam*HI, *Bgl*II or *Eco*RI (see Figures 13-19). This experiment was done for seven human tRNA gene-containing fragments including λHtM2 (one $tRNA^{Tyr}$ gene), λHtM4 (two $tRNA^{Tyr}$ genes, one $tRNA^{Ala}$ gene), λHtM6 (four $tRNA^{Tyr}$ genes), λHt4 (one $tRNA^{Lys}$ gene, one $tRNA^{Gln}$ gene, one $tRNA^{Leu}$ gene), λHt7 (two $tRNA^{Phe}$ genes, two $tRNA^{Lys}$ genes), λHt8 (one $tRNA^{Gly}$ gene), and λHt9 (one $tRNA^{Gly}$ gene). The exact description of the probe sequences used is contained in the Methods section. In all cases, specific hybridization to human sequences was demonstrated. When hybridization to sequences contained within the mouse genomic DNA was observed, the fragments detected were different in size from the hybridizing fragments observed with the human genomic DNA, thus it was feasible to detect hybridization to human tRNA gene-containing fragments within a human-mouse somatic cell hybrid chromosomal mapping panel. Hybridization of the λHt8 specific probe generated a smear corresponding to all sizes of human DNA which was not seen with the mouse

genomic DNA. It appears that this probe must contain a human specific highly repeated sequence element. Limited attempts to remove the element that caused this hybridization pattern were unsuccessful, thus this probe was not used for subsequent hybridization to the mapping panel. The λHt7 probe hybridized to a ~12 kb fragment of *Bam*HI digested human genomic DNA. The same probe hybridized to a ~4.4 kb fragment in the chromosomal mapping panel, therefore this may be a restriction fragment length polymorphism. The human DNA used for preliminary hybridization studies is from a single unknown individual. It is possible that the DNA from this source may contain a rare polymorphism that is not found in the human DNA used in the construction of the chromosomal mapping panel.

The chromosomal locations of fragments contained within the human-λ recombinants λHtM2 (one $tRNA^{Tyr}$ gene), λHtM6 (four $tRNA^{Tyr}$ genes), λHt7 (two $tRNA^{Phe}$ genes and two $tRNA^{Lys}$ genes), and λHt9 (one $tRNA^{Gly}$ gene) have been determined (see Figures 20-23: Tables 1-3). The determination of chromosomal locations for the other human-λ recombinants was not possible because the membranes used for the mapping experiments developed areas of extreme non-specific background radioactivity that masked hybridization results and appeared to be independent of the probe DNA used for hybridization. Lambda HtM2 and λHtM6 were both shown to be located on chromosome 14. It is interesting that these two independently isolated human-λ recombinants were both shown to be located on the same chromosome. Moreover, it has been demonstrated that these recombinants in fact contain overlapping regions of the human genome (D. Spadafora-personal communication). This result, and the observation of a second human $tRNA^{Tyr}$ gene as well as a $tRNA^{Ala}$ gene on λHtM4 (this study; D. Spadafora-personal communication), suggests that the members of this tRNA gene family are clustered and may be contained on one or a few human chromosomes. Other examples of tRNA genes mapping to this chromosome have been described. A tRNA gene heterocluster containing single genes for $tRNA^{Leu}$ and

$tRNA^{Thr}$ plus two genes for $tRNA^{Pro}$ (Olsen et al., 1987) and sequences homologous to two $tRNA^{Gly}$ loci from chromosome 16 have been shown to be located on this chromosome (McBride et al., 1989).

The human DNA segment from λHt9 was determined to be located on chromosome 1. This is the second example of a glycine tRNA gene being mapped to chromosome 1, in addition to that reported by McBride et al., (1989). A single glutamic acid tRNA gene and gene-like sequences were also shown to be located on chromosome 1 (Boyd et al., 1989). Additionally, most if not all of the gene copies for $tRNA^{Asn}$ have been proposed to reside on this chromosome (Buckland, 1989). Two other $tRNA^{Gly}$ genes have been located on chromosome 16, as well as those homologous sequences already mentioned that have been found on chromosome 14. Another human tRNA gene heterocluster containing a $tRNA^{Gly}$ gene was located on chromosome 17 (Morrison et al., 1991). This $tRNA^{Gly}$ gene was subsequently proven to be the same sequence contained within λHt8 (Doran et al., 1988) and this segment of DNA was also shown to be syntenic with the human tRNA genes contained within λHt4 (Roy et al., 1982). Thus, the human sequences from λHt8 and λHt4 have been demonstrated to reside on chromosome 17. For the human $tRNA^{Gly}$ gene family, it would appear that these genes are dispersed in the human genome since most human $tRNA^{Gly}$ genes have been isolated as orphan tRNA genes and these are found on at least four different human chromosomes.

The human DNA fragment contained within λHt7 was located on chromosome 11. This fragment of DNA contains two human $tRNA^{Phe}$ and two $tRNA^{Lys}$ genes. This is the first example of human tRNA genes having been detected on this chromosome. This brings the total number of human chromosomes shown to contain tRNA genes (or pseudogenes) to nine. Furthermore, this is the first demonstration of the chromosomal location for a human $tRNA^{Phe}$ gene(s). Another $tRNA^{Lys}$ gene contained on the human DNA fragment within λHt4 was previously shown to be

located on chromosome 17 (Morrison et al., 1991). Thus, a limited number of studies of human tRNA gene loci have shown that tRNA genes are dispersed on more than one third of the human chromosomes. It would appear that as studies in tRNA gene localization progress, these genes will probably be found dispersed on most, if not all, of the chromosomes of the human genome. The dispersal of human tRNA genes throughout the genome, combined with the biological importance of these genes, suggests that these loci may serve as important markers for sequence tagged sites (STSs) for human genome mapping studies.

The duplication and dispersal of functional tRNA genes has several important biological consequences. As it has been observed for the differentially regulated 5S rRNA gene family (Darby et al., 1988) and the amplified rRNA genes in *Xenopus* oocytes, an increased rate of transcription is possible due to multiple initiation events from duplicate identical tRNA genes. Multiple gene copies also prevent the disrupting effects that might be associated with a mutation in a single copy tRNA gene. The dispersion of multiple copies of each tRNA gene could provide diversity of 5'-flanking sequences, which are known to modulate the expression of some human tRNA genes (reviewed in Sharp et al., 1985; Pratt et al., 1985; Doran et al., 1987; Arnold and Gross, 1987; Capone, 1988; Shortridge et al., 1989). The occurrence of a number of members of an isoacceptor tRNA gene family with variable extragenic control regions might provide a basis for tissue-specific or differentiation-specific expression of tRNA isoacceptors (Lin and Agris, 1980; Hatfield and Rice, 1986). However, any specific role for tRNA gene duplication and dispersal at this point is conjecture.

4.3 Isolation of human tRNA Genes

In order to understand further the structure, genomic organization and transcriptional regulation of human tRNA genes, it is necessary to acquire more examples of genes coding for a given tRNA gene family. There are estimated to be

more 1000 tRNA genes per haploid human genome, and of these, no more than a few examples have been characterized for any one gene family. For a few examples, such as $tRNA^{His}$, $tRNA^{Trp}$, and $tRNA^{Cys}$, no human genes coding for these tRNAs have been isolated. The polymerase chain reaction represents a powerful tool for the isolation of tRNA genes directly from the human genome. By utilizing various of the following PCR strategies that are based on general information known about human and other mammalian tRNA structure and gene organization it is possible to isolate numerous novel examples of human tRNA genes.

4.3.1 tRNA PCR

The first PCR strategy examined is termed "tRNA PCR". This technique utilizes primers that are based on consensus sequences for various tRNA gene coding sequences and capitalizes on the fortuitous arrangement of human tRNA genes in clusters. This technique had originally been described for intron cloning of human $tRNA^{Tyr}$ genes (Green et al., 1990). Previously it had been demonstrated that human $tRNA^{Tyr}$ genes are arranged in clusters (MacPherson and Roy, 1986; MacPherson, 1988). This made this tRNA gene family the obvious choice for study using tRNA PCR. Amplification products from PCR using both the 5' and 3' gene-specific primers for this gene family were cloned and sequenced. Four fragments of 709, 711, 712, and 715 bp and one fragment of 659 bp in length were isolated, in addition to the expected tRNA gene coding sequences (~95 bp). The larger fragments were subsequently shown to contain two human $tRNA^{Tyr}$ genes arranged as direct tandem repeats. Surprisingly, the flanking sequence between the two genes in each of these larger fragments was shown to be greater than 98% homologous. In fact, most of the diversity of these sequences was found in the intron sequences of the tRNA genes contained at the fragment termini. One of the fragments, pHy20, is probably the same as a human DNA fragment (pJM4) previously isolated by MacPherson (1988) and

further characterized by D. Spadafora (unpublished). Extended sequence analysis of this latter fragment, which was isolated from a human-λ recombinant library, has confirmed the arrangement of human $tRNA^{Tyr}$ genes as tandem repeats and has shown that the sequences of pHy20 and part of pJM4 are greater than 99% homologous, including each of the intron sequences (D. Spadafora - personal communication). A total of 7 bp differences was found between pJM4 and pHy20. Five of these differences are clearly artifacts in the pHy20 sequence that represent a rearrangement in the primer binding site, either during the PCR or during the cloning step. The other 2 bp differences occur in the flanking sequences. One of these represents a polymorphism in the putative termination signal flanking the upstream $tRNA^{Tyr}$ gene. A string of 8 consecutive T nucleotides has been described in pJM4 (MacPherson, 1988) whereas the proposed terminator sequence in pHy20 consists of only 7 T residues. This polymorphic form was previously described by van Tol et al. (1987). The remaining polymorphism has not been previously described. It is possible that this reflects a sequencing error. Alternatively, the error rate for *Taq* DNA polymerase has been estimated to be $\sim 10^{-4}$ (Tindall and Kunkel, 1988). At this rate, fragments of the size characterized in this study might be expected to contain a single incorporation error.

It is probable that some of the other fragments isolated that contained different $tRNA^{Tyr}$ genes at the ends of these DNA segments are artifacts. Although the concentration of primers utilized in this experiment was in an extreme molar excess relative to the concentration of amplified tRNA gene coding sequences produced during the course of the PCR, it is possible that the tRNA gene sequences amplified would also serve as primers for amplification from the template sequence characterized within pHy20. In this manner it would be possible to 'cassette' alternative $tRNA^{Tyr}$ gene-coding sequences with different intron sequences onto the original legitimate fragment.

Figure 49 - Schematic representation for the generation of artifact sequences by tRNA PCR. The top figure represents the fragment characterized for pHy20, which contains two tyrosine tRNA genes arranged as tandem repeats. After denaturation, other amplified tyrosine tRNA genes, which contain different introns (shaded areas), anneal to the longer fragment and serve as primers for polymerase extension.

To discern whether these other fragments contained in pHy6, pHy10, and pHy15 actually represent sequences contained within the human genome, PCR was performed using different primers that were specific for each of the intron sequences and a second primer that was contained within the flanking sequence of the tRNA genes and was common to each of the sequences characterized. The different intron-specific primers used for these PCRs were unique sequences, which when used in combination with an intergenic primer would allow differential amplification. PCR using these primers was performed on the recombinant plasmid DNA and also on human genomic DNA. As was expected, amplification using the intron primer specific for pHy20 was observed when both the plasmid DNA and human genomic DNA was used as the template. PCR using the intron primers for pHy10 and for pHy15 resulted in amplification when the plasmid DNA was used as the template but not when the template was the genomic DNA. Thus, these sequences are probably artifact sequences, most likely generated by the amplified tRNA gene sequences themselves serving as primers for amplification from another DNA fragment. Surprisingly, however, PCR using the intron primer for pHy6 generated amplification products of the expected size from both the plasmid DNA and from the genomic DNA samples. Thus, it would appear that this second fragment containing two different $tRNA^{Tyr}$ genes may also be contained within the human genome. Twenty two nucleotide differences were observed in the intron sequences for the two tRNA gene sequences contained on pHy6 and pHy20. An additional three differences were observed in the

intergenic flanking regions. Due to the number of sequence differences between the human DNA fragments from pHy6 and pHy20 these sequences are probably not alleles. Hybridization of the *Apa*I-*Hin*dIII fragment from pJM4 (=pHy20/pHy6) to *Bam*HI, *Bgl*II or *Eco*RI digested human genomic DNA resulted in only a single hybridizing band in all cases. This would suggest that both of these sequences are contained within a single region of the human genome.

These results can be explained by a tandem duplication mechanism for gene dispersal for these $tRNA^{Tyr}$ genes. The high degree of homology between pHy20 and pHy6 also suggests that these sequences were recently duplicated. However, no known mechanism has been proposed that would allow conservation of the tRNA gene exon and intergenic flanking regions while allowing divergence of the intron sequences. Other human $tRNA^{Tyr}$ genes have been found to occur in clusters (MacPherson, 1988), and the genes for $tRNA^{Tyr}$ in *X. laevis* are part of a tandemly repeated tRNA gene cluster (Müller and Clarkson, 1980). An analysis of the flanking sequence from pJM4 revealed that there were no direct repeats in the flanking regions as might be expected if these genes were dispersed by a retroposition mechanism. Observations regarding a tandemly duplicated $tRNA^{Gly}$ gene locus provide evidence for both models of gene duplication (McBride et al., 1989). One of these glycine tRNA genes exhibited identical 8 bp direct repeats in the flanking regions, supporting a retroposition mediated duplication, while the second glycine tRNA appears to be the result of a subsequent tandem reduplication event.

A number of novel $tRNA^{Tyr}$ intron sequences was also characterized by tRNA PCR. Five previous examples of introns for this gene family were shown to be 20 or 21 bp in length (MacPherson, 1988). It was expected that longer intron sequences might be isolated since the introns for tRNA genes in other species are often considerably longer. However, a 16 bp intron was observed for a human $tRNA^{Tyr}$ pseudogene (van Tol et al., 1987). The finding of an similar intron sequence of 16 bp

in length for a human $tRNA^{Tyr}$ gene was first demonstrated by Green et al. (1990) using this PCR technique. Additionally, in that study intron sequences were also described for human $tRNA^{Leu}$ genes. Previous examples of human $tRNA^{Leu}$ genes (Roy et al., 1982; Chang et al., 1986) that had been isolated from a human genomic DNA library showed no intron sequences. The intron sequences in human $tRNA^{Tyr}$ genes were shown to be a prerequisite for pseudouridine biosynthesis in the anticodon (van Tol and Beier, 1988). Like the $tRNA^{Tyr}$ genes, the human $tRNA^{Leu}$ genes that contain introns also contain a modified nucleotide in the anticodon. Therefore, the presence of introns may be required for correct processing of these tRNAs. Differences in intron sequences or the absence of introns probably reflects the absence of selective pressure for strict sequence conservation.

Alternatively, it might be possible that tRNA gene sequence differences (e.g. introns) represent a transition state in the evolution of these genes. The absence of introns in some human $tRNA^{Leu}$ genes may be evidence for retropositional tRNA gene duplication. The RNA intermediate involved in a RNA-mediated transposition can be either a processed RNA polymerase II or a RNA polymerase III transcript which lacks intron sequences (Weiner et al., 1986; Rogers, 1985, Vanin, 1985). RNA polymerase III transcription of tRNA genes normally initiates 3-15 nucleotides upstream from the coding sequence and terminates at a stretch of at least four T nucleotides at variable positions downstream from the gene (Ciliberto et al., 1983; Sharp et al., 1985; Geiduschek and Tocchini-Valentini, 1988). Processing of transcribed tRNAs involves sequential removal of the 5'-leader and 3'-trailer sequences and removal of intron sequences, if present (Sharp et al., 1985; Castano et al., 1985, 1986; Johnson and Abelson, 1983; van Tol and Beier, 1988). This may explain the absence of introns in most human tRNA gene families and the absence of introns in some human $tRNA^{Leu}$ genes, depending upon the nature of the RNA intermediate involved in the transposition event. Moreover, the presence of introns in all known examples of human $tRNA^{Tyr}$

genes may mean that these genes were duplicated by tandem duplication thereby preventing intron elimination during gene dispersal, if a processed RNA intermediate was involved. The absence of a 3' terminal CCA in human tRNA genes (added post-transcriptionally in eukaryotes), however, represents a weakness in the RNA mediated retroposition mechanism for tRNA gene dispersal, again depending upon the structure of the intermediate.

The results herein for tRNA PCR using tyrosine tRNA specific primers were not the same as those reported by Green et al., (1990), in which amplification of the larger fragments was not observed. The differences in results between these two sets of experiments may be due to the different primer sequences used for PCR, or to differences in the PCR conditions. The primers used in this study each contained 21 nt of sequence that was complementary to the sequence contained within the human $tRNA^{Tyr}$ gene, whereas the primers utilized by Green et al, (1990) possessed 15 nt of complementary sequence with additional oligonucleotide sequence comprised of linker sequence that was included for subsequent cloning steps. Perhaps a more important factor may have been the length of time that was allowed for the extension stage of the PCR. The extension stage of PCR for this study was 3 min compared to a 1 min extension stage in the other study. This extra time may be critical for allowing complete synthesis of longer sequences, especially when this synthesis is competing with the amplification of smaller fragments that would be expected to occur more efficiently. In this study, PCR was performed for 30 cycles compared to 25 cycles. The extra cycles may be required for the synthesis of longer amplification products under the conditions of these experiments. One other difference was the buffer composition used for each study, although the effect of this variable is unclear.

In order to demonstrate the general applicability of this technique for the amplification of human tRNA gene-containing fragments, tRNA PCR was then performed on human genomic DNA using primers specific for the consensus sequence

of human $tRNA^{Phe}$ coding sequences. This gene family was chosen because it was known that $tRNA^{Phe}$ genes could be clustered since two $tRNA^{Phe}$ genes have been found arranged as tandem repeats separated by ~9 kb within the human DNA segment contained in λHt7 (Doran et al., 1987). The amplification products from single primer PCR utilizing either the 5' or 3' gene-specific primer individually, or from PCR using both primers were cloned and sequenced. A 1.1 kb fragment resulting from amplification using the 5' gene-specific primer alone hybridized to the 3' gene-specific primer, thus this fragment was judged likely to contain at least one complete $tRNA^{Phe}$ sequence. This clone was designated pHf1-39. Sequencing of this fragment showed that this DNA segment was generated by amplification between a single $tRNA^{Phe}$ gene and a pseudogene arranged as inverted repeats relative to one another. The coding sequence for the gene was shown to be the same as previous examples of human $tRNA^{Phe}$ genes. A possible termination sequence was found downstream from the tRNA gene. A smaller ~0.8 kb fragment did not hybridize to this probe and was not characterized further.

The products of amplification using only the 3' gene-specific primer individually in a PCR were cloned. Two different fragments of ~450 and 500 bp in length were sequenced. The larger fragment (pHf7-31) was the result of amplification between a $tRNA^{Phe}$ gene and a tRNA-like sequence which contained only the primer binding site (completely lacking other sequences from the $tRNA^{Phe}$ gene coding sequence). The 5' flanking sequence of this gene was shown to be unlike the sequences found in pHf1-39 or within $tRNA^{Phe}$ sequences previously characterized (Doran et al., 1987). The second fragment (pHf7-30) was shown to be the result of amplification between two tRNA-like sequences that contained only the sequence of the primer binding site at one terminus and a putative pseudogene at the other terminus. A reiterated sequence with a core consensus sequence of 'GAXT' was found tandemly repeated 13x in this fragment. In addition to the amplification of tRNA gene

sequences, only the 492 bp fragment, containing a complete $tRNA^{Phe}$ gene at one terminus, was observed when both primers were included in the PCR. The longer fragment (pHf1-39) derived from amplification using the 5' gene-specific primer alone was not observed. Perhaps the additional competition for limiting substrate(s) that might be expected to occur when multiple amplification events are ongoing may be responsible for preventing synthesis of longer fragments. This competitive effect is not clearly understood, however. The amplification of sequences contained within pHf7-31 might be outcompeted by simultaneous amplification of the $tRNA^{Phe}$ gene sequences, but this was not observed. Amplified tRNA genes might also serve as primers, thus lessening the effect of competition for primers experienced for longer sequences containing complete tRNA genes. In other PCRs for $tRNA^{Arg}$, $tRNA^{Met}$, $tRNA^{Lys}$, and $tRNA^{Trp}$, none of the amplification products from one primer PCR was observed when both primers were included in a PCR.

This type of experiment was repeated using primers that were specific for the human $tRNA^{Lys}$, $tRNA^{Arg}$, $tRNA^{Met}$, and $tRNA^{Trp}$ gene families. No amplification was observed when the $tRNA^{Lys}$ 5' gene-specific primer was used individually in a PCR. Amplification from the 3' gene-specific primer resulted in the production of 2 very weakly hybridizing fragments of approximately 1 and 4 kb in size. When both primers were used in a PCR the only amplification observed was that of the $tRNA^{Lys}$ gene sequence.

A strongly hybridizing fragment of ~1.7 kb in length and two weakly hybridizing bands of ~0.25 and 0.35 kb were detected by hybridizing an arginine tRNA 5' gene-specific primer to PCR products resulting from amplification using the same $tRNA^{Arg}$ primer. None of these fragments hybridized to the 3' gene-specific primer however. Conversely, three fragments of ~1.3, 0.80, and 0.55 kb in length that were derived from PCR using the 3' gene-specific primer were shown to hybridize to the 5' gene-specific primer. These fragments may contain complete $tRNA^{Arg}$ gene sequences.

Two smaller fragments of ~0.40 and 0.25 kb in length did not hybridize to the 5' gene-specific primer and are thus probably devoid of complete $tRNA^{Arg}$ gene coding sequences.

Numerous amplification products were detected in single primer PCR using the $tRNA^{Met}$ 5' gene-specific primer, the $tRNA^{Trp}$ 5' gene-specific primer or the $tRNA^{Trp}$ 3' gene-specific primer. Each of the fragments produced in these PCRs, however, hybridized only with the probe which was specific for the primer that was used in the PCR. Because none of the amplification products in these cases hybridized with the probe specific for the other terminus of the tRNA gene sequence it is unlikely that these amplification products contain active tRNA genes but were rather the result of amplification between tRNA pseudogenes.

These results suggest that while clustering of identical tRNA genes may be relatively infrequent (i.e. only a few amplified bands which contain two tRNA genes for any gene family), amplification between a tRNA gene and a pseudogene or between pseudogenes is likely to occur for most human tRNA gene families. Many examples of tRNA pseudogenes that have been characterized to date consist of only slight mutations that disrupt the ability of a tRNA transcript to assume the final correct tertiary structure that is necessary for function. Other pseudogenes, however, have been shown to contain relatively large changes to the tRNA gene sequence. A recombinant plasmid containing *Drosophila melanogaster* DNA was shown to contain a region of homology to the correct gene sequence for $tRNA_i^{Met}$ only for the coordinates from 7 through 39. Hybridization of this plasmid to genomic DNA demonstrated that there are approximately 30 regions of homology contained within this genome. The hybridization of a larger fragment encompassing this segment of DNA displayed hybridization kinetics of middle-repetitive DNA (Gergen et al., 1981). It was proposed that this DNA segment represented a 'fragmented' initiator tRNA gene that may have been brought about by repeated insertion and excision of a transposable element(s) into

the tRNA gene coding region. A number of pseudogenes have been characterized in a rat tRNA gene heterocluster encoding genes for $tRNA^{Asp}$, $tRNA^{Gly}$, and $tRNA^{Glu}$. This cluster of tRNA genes is reiterated approximately 10 times in the haploid genome. Changes to tRNA gene sequences include deletions of up to 14 nucleotides in length (Shibuya et al., 1982) including a 11 nt region from the 3' end of a $tRNA^{Glu}$ gene. Similarly, a mouse $tRNA^{Phe}$ pseudogene was characterized that was shown to lack the coding sequence from coordinates 1 to 38 (Reilly et al., 1982). The pseudogene in this case included a 3' terminal CCA, which is normally not found in eukaryotic tRNA gene coding sequences but is added post-transcriptionally. This pseudogene may be the result of the information from a mature tRNA that was reverse transcribed and integrated into the mouse genome (Reilly et al., 1982).

4.3.2 Alu-tRNA PCR

There is evidence suggesting that many mammalian short interspersed repetitive DNA sequence (SINE) families represent amplified tRNA retropseudogenes (Weiner et al.,1986; Rogers, 1985; Daniels and Deininger, 1985). Additionally, it has been suggested that retroposons have a tendency to integrate into the tails of other retroposons (Rogers, 1985; Weiner et al., 1986). The observed preponderance of Alu repeat sequences in the flanking regions of human tRNA genes may be the result of such a phenomenon. The role of these repeat sequences, as well as any role of tRNA pseudogenes in tRNA gene flanking regions is unknown. It is possible that these sequences are involved in binding transcription factors, or are simply traces of the evolutionary history of tRNA genes.

In order to study further the nature of the association between human tRNA genes and the Alu repeat sequence family, PCR was performed using primers specific for the 5' conserved regions of the human $tRNA^{Lys}$ and $tRNA^{Phe}$ gene families and primers specific for different regions within the consensus sequence for the human Alu

repeat sequence family. Because the most likely amplification event in such a PCR would be that which occurs between two Alu sequences, a two stage PCR was designed that allows for increased selection of amplification products containing tRNA genes. When PCR was performed using the $tRNA^{Lys}$ 5' biotinylated gene-specific primer and four different Alu primers, amplification products of various lengths were produced. This production of DNA fragments of heterogeneous lengths was also seen when the Alu specific primers were used individually in a PCR. The PCR was repeated, using as template DNA those amplification products of the first Alu-tRNA PCR that had been selected using streptavidin agarose for $tRNA^{Lys}$ gene-containing sequences. For each of these reactions the amplification of discrete fragments was observed. These fragments ranged in length from ~0.2 to 0.7 kb. Similar results were seen for two of the four Alu primers when used in a PCR with the 5' $tRNA^{Phe}$ gene-specific primer. For both tRNA gene families the best results were obtained for the PCR using the same Alu specific primer (Alu2). This result may suggest that this primer more closely resembles the actual sequence of the Alu repeat family members that actually flank some tRNA genes than the Alu primers that were used in the other PCRs. Alternatively, it may be that a particular orientation of Alu repeat sequence and human tRNA genes is the most predominant, at least within an amplifiable distance.

The sizes of the amplification products for many of the PCRs for both tRNA gene families suggest that the tRNA gene sequences and the Alu repeat sequences are overlapping. This type of arrangement has been observed previously. There is evidence that a human opal suppressor pseudogene (O'Neill et al., 1985) was formed by retroposition. This sequence is located on human chromosome 22, while the functional gene (described by Leinfelder et al., 1988, as a selenocysteine tRNA gene) is found on chromosome 19 (McBride et al., 1987). This putative retropseudogene appears to have been inserted into the junction of two Alu sequences fused in a head-to-tail fashion, and this insertion is associated with a 17 bp 3'-truncation and 11-bp 5'

truncation of the two adjacent Alu sequences. Another arrangement of tRNA gene and Alu sequence has also been described in the case of a valine tRNA gene that was found within a human tRNA gene heterocluster and was shown to be directly adjacent to an Alu sequence (McBride et al., 1987) oriented in a tail-to-tail fashion.

4.3.3 Inverted PCR

A third strategy that was used to isolate human tRNA genes directly from the human genome is "Inverted PCR". The advantage of this method, in principle, is that it should be possible to isolate novel flanking sequences from both the 5' and 3' regions of any tRNA gene by using primers that are complementary to the consensus sequence for the tRNA gene itself. This method should prove very useful in the study of human tRNA genes, especially for transcriptional regulation studies and genome mapping studies, where more extensive knowledge of flanking sequences is necessary.

Inverted PCR was performed using primers specific for the human $tRNA^{Phe}$ gene family. Human genomic DNA was digested with *Cfo*I, *Hae*II, or *Hin*PI. The digested DNA was diluted serially tenfold (from 10^0 to 10^{-5}) for use in a ligation reaction to allow the circularization of this DNA, and the resulting templates were subjected to PCR. After size fractionation of amplification products and ethidium bromide staining of the agarose gel, the greatest number of fragments was observed for the 10^0 dilution . Weaker amplification also occurred for the 10^{-1}, 10^{-2}, and 10^{-3} dilutions (data not shown). The weaker results for more highly diluted samples probably represent the dilution of suitable template molecules below a critical level required for detection by PCR (using the parameters in this study) rather than an effect on circularization efficiency. Similar fragments appear to have been amplified in reactions using the same primers on template DNA that had been digested with different restriction enzymes. However, those fragments which hybridized most efficiently to a probe complementary to the $tRNA^{Phe}$ gene coding sequence (that is not found in either

of the PCR primer sequences) were not the most intensely stained fragments. It is possible that amplification had occurred directly from the genomic DNA between tRNA gene sequences or gene-like sequences that are found clustered, and this non-inverted PCR was responsible for some or all of the fragments generated. Therefore, inverted PCR may have generated only some of the fragments. When inverted PCR was performed on human genomic DNA digested with *Hin*dIII, the best amplification occurred with the undiluted sample, however weak hybridization results were seen with the sample that had been diluted 10^{-5} fold prior to ligation and PCR. Differences in the apparent results (as determined by hybridization) may be due to differences in the starting mass of DNA used or differences in the efficiency of restriction endonuclease digestion. Examination of an aliquot of the human DNA that was digested with *Cfo*I indicated that only partial digestion had occurred and that most of the genomic DNA after digestion was still very large. The effective concentration of DNA ends suitable for ligation/cirularization must have been considerably reduced from that expected for an equivalent mass of human genomic DNA that was completely digested.

The amplification products from inverted PCR for the $tRNA^{Phe}$ gene family were cloned for sequence analysis. A total of 11 different fragments were characterized and all were shown to contain regions complementary to the $tRNA^{Phe}$ 3'-gene probe that was used for screening. Unexpectedly, many of the fragments that were characterized were shown to contain tandemly repeated $tRNA^{Phe}$ genes. As the extent of tandem duplication that was observed in these clones had not been previously described in humans, the possibility arose that many (or all) of these fragments were artifacts generated during the PCR process. Sequence analysis of the fragments contained within pHae2, pHae19, pHae20 and pHae23 revealed that each intergenic flanking sequence for DNA segments containing multiple tRNA genes was also identical and was tandemly repeated for those clones derived from PCRs using the same template DNA. Moreover, the smaller fragments containing 0, 1, or 2 internal

complete $tRNA^{Phe}$ gene sequences are identical to part of the sequence of a larger clone containing 3 internal $tRNA^{Phe}$ gene sequences. The smallest of these fragments analyzed was shown to represent the basic repeated unit consisting of usually ~90 bp DNA segment that contains the 5' tRNA gene-specific primer binding site, an intergenic flanking region of ~40 bp and the 3' end of the $tRNA^{Phe}$ gene coding sequence containing the 3' primer binding site.

Many fragments failed to exhibit the restriction recognition site that would be expected if inverted PCR was the mechanism by which these fragments were amplified. PCR performed on undigested/unligated human genomic DNA resulted in the production of a similar pattern of bands visualized by ethidium bromide staining as seen for inverted PCR using the same primers. However, there was absolutely no hybridization of the 3' $tRNA^{Phe}$ gene-specific probe to the amplification products in this case (data not shown). It would appear, based on these results, that the tandemly repeated organization of $tRNA^{Phe}$ genes found in these clones is not found in the human genome. Incomplete digestion by the restriction endonuclease used in the first step would not explain the generation of the artifact sequences. A clear reconciliation of these results is difficult. The terminal 16 bp sequence from the 3' end of the tRNA gene must be template encoded since this is not found in the primer sequences. Additionally, the segment of DNA between the 5' ends of each of the inverted PCR primers (~9 bp) must also be template encoded. Thus, any explanation invoking primer artifacts cannot easily resolve the observed results. One explanation for the exact tandem repeats of tRNA gene plus flanking sequence might be that PCR had occurred from a circularized template containing a single $tRNA^{Phe}$ gene. For each cycle of PCR, synthesis proceeded for multiple rounds around the circular template thus duplicating the tRNA gene sequence plus flanking sequence (see Figure 50). Circular permutation of the basic unit sequence could be avoided by performing a second restriction endonuclease digestion after the ligation step, using an enzyme which cuts in the tRNA

gene sequence between the primer binding sites (e.g. *Bgl*II for $tRNA^{Phe}$). However, digestion with this second restriction endonuclease would also eliminate amplification of tandemly repeated tRNA genes, if present. The absence of the expected restriction endonuclease recognition site used in the original digestion suggests that the template DNA was not circular and, therefore, weakens this explanation.

Sequence analysis of two other clones (pCfo2 - 145 bp and pCfo35 - 165 bp) revealed a region of exact homology between the two fragments except for the deletion of approximately 20 bp in pCfo2 immediately flanking the 5' $tRNA^{Phe}$ sequence. The sequence for pCfo2 is also devoid of the recognition site of the restriction endonuclease (*Cfo*I) used in the initial stages. A third recombinant, designated pCfo1, was sequenced and shown to contain an insert of 244 bp in length. This fragment contains a single complete internal $tRNA^{Phe}$ gene with identical 5' and 3' intergenic flanking sequences. The flanking sequence was also shown not to contain the restriction endonuclease recognition site. There appears to very little flanking sequence homology between the sequences contained in pCfo1 and those of pCfo2 or pCfo35. None of these sequences showed flanking sequence homology to the fragments contained in pHae2, pHae19, pHae20 or pHae23.

A number of artifacts, generated either during the inverted PCR process or during the cloning of amplification products, were characterized. This includes "primer dimer" where the IR-Phe primer was dimerized, unexpectedly, in a head to tail fashion (data not shown). This primer dimer sequence was found incorporated into a larger fragment (see Figure 40B). Sequence analysis of pCfo35, revealed that *Taq* DNA polymerase had added "TATACGTATATA" to the end of the fragment, possibly in a

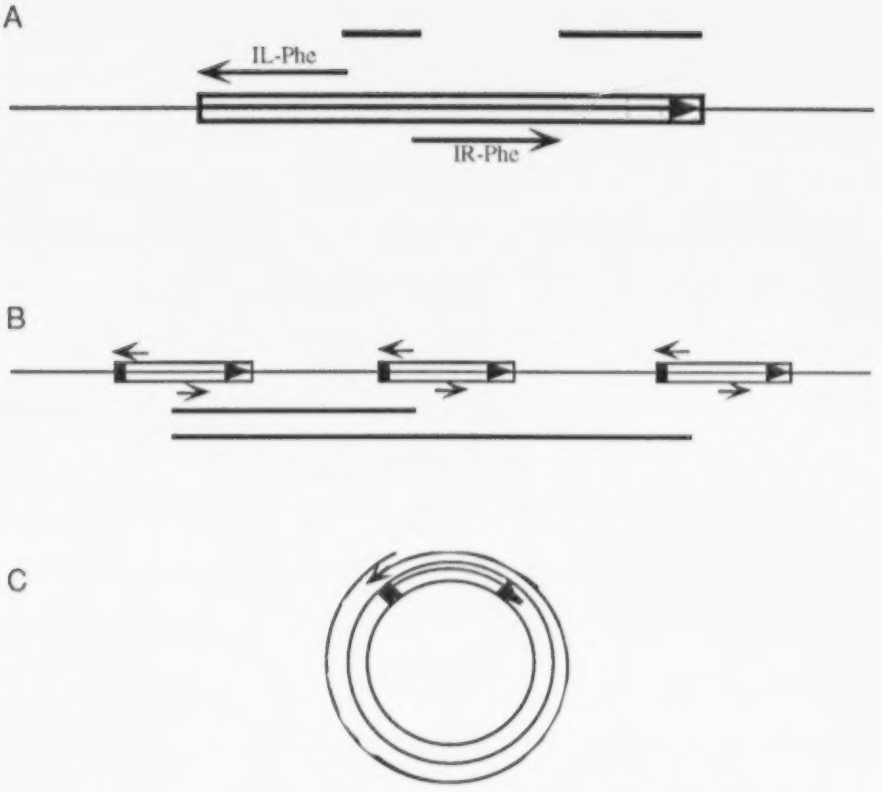


Figure 50 - Schematic representation for the generation of "artifacts" from inverted PCR. The arrangement of PCR primers (arrows) relative to a tRNA gene is shown in (A). The solid bars above the diagram indicate regions in the tRNA gene that are not found in the primer sequences but have been found in amplification products from inverted PCR for $tRNA^{Phe}$. Diagram B shows a tandem duplication of a tRNA gene and some potential amplification products (solid bars) for PCR using the primers shown in A. The intergenic flanking sequences represented here do not contain the restriction endonuclease recognition site. Diagram C shows a circularized template, containing a tRNA gene. The circular arrow represents a primer extension continuing multiple times around the template, generating exact repeats of gene plus flanking sequence.

non-template dependent manner (data not shown). The sequence of pHnp11 (Figure 40D) showed that the IL-Phe primer binding site was inverted. In this orientation, amplification from this sequence should not be possible. Numerous other sequences showed various degrees of truncation of the termini (see Figure 39 and Figure 40C). These latter two artifacts (sequence truncation and primer inversion) were probably generated during the cloning step, since the sequences characterized are not suitable

templates for inverted PCR. Ligation of PCR generated fragments which have experienced exonucleolytic degradation may be selected for during cloning since these fragments would have a 5' PO_4. Other potential artifacts, including the circular permutation of sequences (Figure 39) and the lack of restriction endonuclease recognition sites (Figure 38-40) have already been described.

Four different clones were recovered from the amplification products of inverted PCR performed on human genomic DNA that was initially digested with *Hin*PI prior to ligation and PCR. The fragments contained within pHnp2, pHnp4 and pHnp11, and pHnp12 are approximately 1100 bp, 400 bp, 700 bp and 240 bp in length, respectively. Partial sequence analysis of the fragments in these clones revealed that each contains the restriction recognition site for *Hin*PI. The complete sequence of pHnp12 was determined. This fragment was shown to have the 5' and 3' $tRNA^{Phe}$ gene-specific sequences at the fragment termini and was also shown to contain a single internal $tRNA^{Phe}$ gene. However, unlike the sequences previously described (e.g. pCfo and pHae) this fragment contains divergent flanking sequences in the 5' and 3' flanking regions of this gene. Primers were designed for $tRNA^{Phe}$ gene flanking regions that were also flanking the *Hin*PI site in non-overlapping directions relative to the sequence contained in the plasmid. In an attempt to reconstitute the correctly oriented, contiguous sequence, PCR was then performed on human genomic DNA using these primers. Ethidium bromide staining of the amplification products revealed a strong fragment of ~800 bp in length. This result is unexpected since a fragment of ~250 bp length should have been produced. Nonetheless, this larger amplification product was shown to hybridize to an oligonucleotide probe that was specific for the 3' terminus of the $tRNA^{Phe}$ gene coding sequence, thus this fragment may contain at least one $tRNA^{Phe}$ gene. The amplification of a single fragment (probably containing a $tRNA^{Phe}$ gene) directly from the human genome using primers that were designed to be specific for a unique region of the human genome and that were oriented in non-

overlapping directions, suggests the sequence contained within pHnp12 was produced by inverted PCR. Despite problems with unexplained amplification products from inverted PCR, the usefulness of this technique for the isolation of tRNA gene-containing fragments has been demonstrated.

The results that were seen for inverted PCR performed for the $tRNA^{His}$ and $tRNA^{Trp}$ gene families similarly demonstrated this dichotomy between ethidium bromide staining of amplification products and subsequent results for hybridization of the 3' gene-specific oligonucleotide probes to the immobilized DNA. Human genomic DNA was digested with *Hpa*II or *Hin*dIII prior to ligation and PCR. PCR was performed and the amplification products separated on a 2% agarose gel. When visualized using ethidium bromide, similar length fragments were observed for each tRNA gene family regardless of which enzyme was used in the restriction digestion step. The hybridization results, however, differed significantly between each of the samples. Thus, it appears likely that multiple events are likely to occur during inverted PCR when using human tRNA gene-specific primers. Because many of the strongly staining fragments do not hybridize to tRNA gene-specific probes, the amplification of these fragments must be due to PCR between incomplete tRNA-like sequences. The number of fragments that are generated in each PCR also suggests that these incomplete tRNA sequences are relatively common. Amplification between 'imperfect' tRNA gene sequences or gene-like sequences, although potentially interesting, could possibly be eliminated by increasing the annealing temperature used in these PCRs. This would limit amplification to fragments which contained nearly perfect matches of the primer binding sites.

The size of the hybridizing amplification products reflects the relative frequency of cutting for each restriction endonuclease, as determined by the recognition sequence of each enzyme. In the histidine tRNA experiment, a single large fragment was amplified from human genomic DNA that had been digested initially with *Hin*dIII

(AAGCTT) compared to the amplification of ~4 smaller fragments produced by PCR from *Hpa*II (CCGG) digested DNA. Because *Hpa*II is expected to cut more frequently in the sequence surrounding a tRNA gene, the average size of amplified fragments should be correspondingly smaller.

4.3.4 Linker PCR

A fourth PCR strategy that was investigated for the amplification of tRNA gene sequences is termed "Linker PCR". This technique involves the ligation of linkers to the ends of digested genomic DNA in order to provide a known 'second site' that would be suitable for PCR. When linker PCR was performed (using a $tRNA^{Phe}$ 5' gene-specific primer plus a linker primer) on digested human genomic DNA that had linkers ligated to the fragment ends, heterogeneous fragments of variable lengths were produced (data not shown). No detectable hybridization was seen when these amplification products were probed. Similar results were seen when only the linker primer was used in a PCR. The majority of amplification products are due to amplification between linker primers that have been ligated to both ends of a variety of DNA fragments. In principle, since the double stranded linkers used in the ligation reaction are non-phosphorylated, ligation should only occur between the 3'-OH of one strand of the linkers and the 5'-PO_4 of the digested DNA. Any single strand of DNA should not have linkers ligated to both ends. Polymerase extension from the linker primer should not result in the synthesis of a second linker primer binding site. One likely explanation, however, is that the *Taq* DNA polymerase simply filled in the complementary sequence of the ligated linker, prior to denaturation of the double stranded template. This problem might be avoided by performing a "hot start" for the PCR, whereby the enzyme is added when the template DNA is completely denatured. Another possibility is that the linker primer was only ligated to one end of the digested DNA, but the other end of the fragment terminated in a hairpin structure. Polymerase

extension using the hairpin structure as a primer would generate a second copy of the linker primer site on the same strand of DNA.

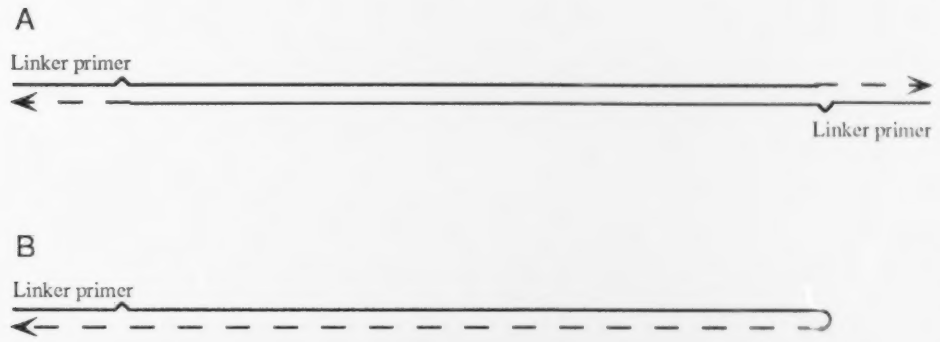


Figure 51 - Schematic representation of the generation of fragments containing linker primer binding sites at both ends. Diagram A shows the enzymatic "filling in" of the linker primer binding site, whereby the DNA polymerase utilizes one strand of the duplex as a primer. Diagram B shows the possible use of a hairpin structure as a primer for the generation of a second linker primer binding site.

In order to overcome this high degree of background amplification, a two stage PCR was utilized in which a first round of PCR was performed using a biotinylated tRNA gene-specific primer and one of the linker primers. The amplification products were then selected for sequences containing the biotinylated primer using streptavidin-agarose. A second stage PCR was then performed using the same combination of primers as used in the original PCR, except that non-biotinylated tRNA gene-specific primers were used. The results of these PCRs again revealed that a high degree of background amplification occurred. However, when the products of these reactions were probed with a complete $tRNA^{Phe}$ gene sequence, strongly hybridizing fragments were detected in those PCRs which used the 3' $tRNA^{Phe}$ gene-specific primer on templates that had originally been digested with *Ssp*I, *Sma*I, *Apa*I, or *Eco*RV. The selection step appears not to have been completely successful. Because a relatively high concentration of biotinylated primers was used in the first round of PCR, any primers which did not initiate amplification products might be expected to compete for a

limited number of binding sites on the streptavidin-agarose. Additionally, a significant amount of contamination with amplification products containing linkers at both ends must have occurred. It is possible that a limited amount of selection resulted, but this was not as efficient as would have been achieved if a lesser concentration of biotinylated primers was used initially and if the selection step was more stringent.

The number of genes coding for any given tRNA has been estimated to be in the range of 10-20 per haploid human genome. Thus, more tRNA gene-containing fragments were expected to be amplified. Non-specific background amplification might have limited the amplification of these tRNA gene-specific sequences. In other experiments where similar PCR techniques were employed, increased specificity of amplification was achieved by using a two stage PCR in which an internal primer was used in the second PCR (Fors et al., 1990). When this was applied to linker PCR utilizing a $tRNA^{Phe}$ gene-specific primer, there was an increase in the intensity of the discrete bands amplified and a corresponding decrease in the background amplification. These products were fractionated on a 2% agarose gel and hybridized to a complete $tRNA^{Phe}$ gene probe prepared by random primer extension. A number of strongly hybridizing fragments were detected in those PCRs which used the internal primer (IR-Phe) specific for the 3' region of the $tRNA^{Phe}$ gene sequence. Other PCRs clearly resulted in the production of strongly staining discrete fragments but these hybridized weakly to the $tRNA^{Phe}$ specific probe. The 5' $tRNA^{Phe}$ gene internal primer (IL-Phe) is specific for the terminal 21 nt of the 5' region of the tRNA gene sequence and is oriented towards the 5' flanking sequence, whereas the 3'-$tRNA^{Phe}$ gene internal primer (IR-Phe) is specific for the anticodon stem and TψC stem region of the tRNA gene and is oriented through the gene towards the 3' flanking sequence. Thus, amplification products from this primer probably contain ~40 nt of tRNA gene sequence. The differences in hybridization intensity may simply reflect the unequal proportion of the tRNA gene sequence contained in the final PCR product and a

correspondingly less stringent hybridization may have yielded better results. In summary, a number of discrete fragments were amplified by this technique and these were shown to hybridize to a $tRNA^{Phe}$ specific probe. It should thus be possible to use linker PCR for the isolation of tRNA gene flanking sequences for a number of gene families. Additionally, it should be possible to apply this method for 'walking' from any specific locus for the study of other DNA sequences.

4.4 Conclusions

A number of probe preparation methods were tested with the intention of examining human chromosomal mapping panels. The results for hybridization of a 19 nt long $tRNA^{Tyr}$ specific oligonucleotide to human genomic DNA revealed that, under the conditions used in these experiments, these probes would not be suitable for gene mapping studies. Similarly, hybridizations using M13 primer extension probes were also found unsuitable because of a high degree of non-specific hybridization. Other conditions might improve the use of these methods for genomic hybridizations, especially with oligonucleotides, but these were not examined. The random primer extension method produced the best results for human genomic hybridizations. When $tRNA^{Gln}$ gene probes containing various amounts of flanking sequence were hybridized to human genomic DNA, strong specific signals were detected. Longer probes, containing more unique flanking sequence, tended to provide better specific hybridization. Human specific hybridization for seven probes, containing various human tRNA genes was demonstrated. The chromosomal locations for four of these probes (λHtM2, λHtM6, λHt7, and λHt9) were determined. Lambda HtM2 and HtM6 both map to chromosome 14. Lambda Ht7 is located on chromosome 11, while λHt9 is found on chromosome 1.

In order to isolate novel tRNA genes for further study, various strategies based on PCR were attempted. By using "tRNA PCR", five fragments containing ten tRNA

genes or pseudogenes were isolated for the $tRNA^{Phe}$ and $tRNA^{Tyr}$ gene families. One fragment (pHy20) is the same as a human DNA segment previously isolated from a human-λ recombinant library. By the use of PCR, a second fragment (pHy6) was shown to be found in the human genome. However, artifactual sequences were shown also to be amplified for the $tRNA^{Tyr}$ gene family because amplified tRNA genes were serving as primers. This was only detectable due to differences in their intron sequences. For most human tRNA genes, which lack introns, this event would not be noticeable. "tRNA PCR" depends on the fortuitous arrangement of tRNA genes (or gene-like sequences) as direct or inverted repeats. As such, this procedure is unlikely to give positive results for all tRNA gene families or all tRNA genes within a family.

Another procedure used to recover tRNA genes was "Inverted PCR". This was done for the $tRNA^{Phe}$, $tRNA^{His}$, and $tRNA^{Trp}$ gene families. At least 11 different fragments containing $tRNA^{Phe}$ genes were cloned and sequenced. Many of these sequences were shown to contain multiple $tRNA^{Phe}$ genes, arranged as tandem repeats. A number of inverted PCR amplification products were probably artifacts. These artifacts, however, may reflect some unknown aspect of the genomic organization of this gene family. Despite the isolation of these sequences, inverted PCR is likely to be very useful for the recovery of novel tRNA gene sequences. To verify that some of the sequences obtained by this method are representative of the actual sequences in the human genome, amplification of $tRNA^{Phe}$ gene-containing sequences from genomic DNA was performed using primers which were based on sequences found in two of these clones. Fragments isolated in this manner are especially interesting since both the 5' and 3' flanking sequence for individual tRNA genes should be present.

Two other PCR strategies ("Alu-tRNA PCR" and "Linker PCR") were attempted in conjunction with streptavidin selection of tRNA gene sequences. By using biotinylated tRNA gene-specific primers with streptavidin selection, it was possible to increase the amount of specific amplification for those PCRs which generate significant

amounts of non-specific background amplification. Human Alu repeat sequences were demonstrated to reside in the flanking regions surrounding some tRNA genes. Both orientations of the Alu sequence, relative to the tRNA gene, were shown to exist. Because Alu repeat sequences are often found surrounding human tRNA genes, "Alu-tRNA PCR" will be moderately useful for the recovery of novel flanking sequences. Despite the high degree of background amplification, "Linker PCR" should prove even more useful, since it will be possible to 'walk' in both directions from a tRNA gene.

In all four PCR strategies it will be important to examine systematically, variations in conditions to find the best, and most stringent parameters for tRNA gene amplification. The number of artifact sequences recovered would probably be reduced by increasing the temperature of the primer annealing during PCR, such that amplification occurs strictly between well conserved tRNA gene sequences. Overall, the use of these PCR strategies, individually or in combination, will allow an increased rate of isolation of novel tRNA genes and their flanking sequences. Recovery and analysis of tRNA gene loci is more efficient than current methods of gene isolation such as screening human-λ recombinant libraries, especially when large number of gene loci are being examined. The time limiting factor in the characterization of amplification products is often associated with cloning PCR products. Novel sequencing strategies, such as direct sequencing, will increase the rate at which information can be obtained and will also improve the results due to the reduction of errors due to cloning of fragments containing incorporation errors. Fragments containing tRNA genes and flanking regions isolated by PCR will be useful for chromosome mapping studies and for experiments determining transcriptional regulation of tRNA gene families.

5. REFERENCES

Arnold, G. J., and Gross, H. J. (1987). Unrelated leader sequences can efficiently promote human tRNA gene transcription. Gene, **41:** 237-246.

Arnold, G. J., Schmutzler, C., Thomann, U., van Tol, H., and Gross, H. J. (1986). The human tRNAVal gene family: organization, nucleotide sequences and homologous transcription of three single-copy genes. Gene, **44:** 287-297.

Baker, R. E. and Hall, B. D. (1984). Structural features of yeast transfer RNA genes which affect transcription factor binding. EMBO J., **3:** 2793-2800.

Bieker, J. J., Martin, P. L., and Roeder, R. G. (1985). Formation of a rate-limiting intermediate in 5S gene transcription. Cell, **40:** 119-127.

Birnboim, H. C. (1983). A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol., **100:** 143-155.

Bonner, T. I., Brenner, D. J., Neufeld, D. R., and Britten, R. J. (1973). Reduction in the rate of DNA reassociation by sequence divergence. J. Mol. Biol., **81:** 123-135.

Boyd, E. Theriaul, A., Goddard, J. P., Kalaitsi, M. Spathas, D. H. and Connor, J. M. (1989). Chromosomal assignment of a glutamic acid transfer RNA ($tRNA^{Glu}$) gene to 1p36. Hum. Genet., **81:** 153.

Braun, B. R., Riggs, D. L., Kassavetis, G. A., and Geiduschek, E. P. (1989). Multiple states of protein-DNA interaction in the assembly of transcription complexes on *Saccharomyces cerevisiae* 5S ribosomal RNA genes. Proc. Natl. Acad. Sci. USA, **86:** 2530-2534.

Britten, R. J., Baron, W. F., Stout, D. B., and Davidson, E. H. (1988). Sources and evolution of human *Alu* repeated sequences. Proc. Natl. Acad. Sci. USA, **85:** 4770-4774.

Brown, D. D. (1984). The role of stable complexes that repress and activate eukaryotic genes. Cell, **37:** 359-365.

Buckland, R. A. (1989). Genomic organization of the human asparagine transfer RNA genes: localization to the U1 RNA gene and class I pseudogene repeat units. Am. J. Hum. Genet., **45:** 283-295.

Burke, D T., Carle, G. F., and Olsen, M. V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science, **236:** 806-812.

Capone, J. P. (1988). Modulation of the phenotypic expression of a human serine tRNA gene by 5'-flanking sequences. DNA, **7:** 459-468.

Carcamo, J., Lobo, S., Merino, A., Buckbinder, L., Weinmann, R., Natarajan, V., and Reinberg, D. (1989). Factors involved in specific transcription by mammalian RNA polymerase II. J. Biol. Chem., **264:** 7704-7714.

Carle, G. F., Frank, M., and Olsen, M. V. (1986). Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. Science, **232:** 65-68.

Castano, J. G., Ornberg, R., Koster, J. G., Tobian, J. A., and Zasloff, M. (1986). Eukaryotic pre-tRNA 5' processing nuclease: Copurification with a complex cylindrical particle. Cell, **46:** 377-387.

Castano, J. G., Tobian, J. A., and Zasloff, M. (1985). Purification and characterization of an endonuclease from *Xenopus laevis* ovaries which accurately processes the 3' terminus of human pre-$tRNA_i^{Met}$ (3' pre-tRNase). J. Biol. Chem., **260:** 9002-9008.

Chalice, J. M., and Segall, J. (1989). Transcription of the 5S rRNA gene of *Saccharomyces cerevisiae* requires a promoter element at +1 and a 14 base pair internal control region. J. Biol. Chem., **264:** 20060-20067.

Chang, Y.-N., Pirtle, I. L., and Pirtle, R. M. (1986). Nucleotide sequence and transcription of a human tRNA gene cluster with four genes. Gene, **48:** 165-174.

Chapeville, F., Lipmann, F., Ehrenstein, G. V., Weisblum, B., Ray, W. J. Jr., and Benzer, S. (1962). On the role of soluble ribonucleic acid in coding for amino acids. Proc. Natl. Acad. Sci. USA, **48:** 1086-1092.

Chen, E. Y., and Seeburg, P. H. (1985). Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA, **4:** 165-170.

Ciliberto, G., Castagnoli, L., Melton, D. A., and Cortese, R. (1982). Promoter of a eukaryotic $tRNA^{Pro}$ gene is composed of three noncontiguous regions. Proc. Natl. Acad. Sci. USA, **79:** 1195-1199.

Ciliberto, G., Castagnoli, L., and Cortese, R. (1983). Transcription by RNA polymerase III. In A. A. Moscona and A. Monroy (Eds.), Current Topics in Developmental Biology (pp. 59 - 88). New York: Academic Press.

Clark, B. F. C. (1978). General features and implications of primary, secondary, and tertiary structure. In Altman, S. (Ed.), Transfer RNA. pp14-17, The MIT Press, Cambridge, Massachusetts.

Collins, F. S., and Weissman, S. M. (1984). Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method. Proc. Natl. Acad. Sci. USA, **81:** 6812-6816.

Craig, L. C., Wang. L. P., Lee, M. M., Pirtle, I. L., and Pirtle, R. M. (1989). A human transfer RNA encoding the major and minor valine transfer RNAs and a lysine transfer RNA., DNA, **8:** 457-471.

Daniels, G. R., and Deininger, P. L. (1985). Repeat sequence families derived from mammalian tRNA genes. Nature, **317:** 819-822.

Darby, M. K., Andrews, M. T., and Brown, D. D. (1988). Transcription complexes that program *Xenopus* 5 S RNA genes are stable *in vivo*. Proc. Natl. Acad. Sci. USA, **85:** 5516-5520.

DeFranco, D., Schmidt, D., and Söll, D. (1980). Two control regions for eukaryotic tRNA gene expression. Proc. Natl. Acad. Sci. USA, **78:** 6657-6661.

Doran, J. L., Bingle, W. H., and Roy, K. L. (1988). Two human genes encoding RNA^{Gly}_{GCC}. Gene, **65:** 329-336.

Doran, J. L., Wei, X., and Roy, K. L. (1987). Analysis of a human gene cluster coding for $tRNA^{Phe}_{GAA}$ and $tRNA^{Lys}_{UUU}$. Gene, **56:** 231-243.

Dugaiczyk, A., Boyer, W. H., and Goodman, H. M. (1975). Ligation of *Eco*RI endonuclease-generated DNA fragments into linear and circular structures. J. Mol. Biol., **96:** 171-184.

Feinberg, A., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem., **132:** 6-13.

Feinberg, A., and Vogelstein, B. (1984). Addendum. Anal. Biochem., **137:** 266-267.

Fors, L., Saavedra, R. A., and Hood, L. (1990). Cloning of the shark Po promoter using a genomic walking technique based on the polymerase chain reaction. Nucl. Acids Res., **18:** 2793-2799.

Galli, G., Hofstetter, H., and Birnstiel, M. L. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature, **294:** 626-631

Geiduschek, E. P., and Tocchini-Valentini, G. P. (1988). Transcription by RNA polymerase III. Ann. Rev. Biochem., **57:** 873-914.

Gergen, J. P., Loewenberg, J. Y., and Wensink, P. C. (1981). $tRNA^{Lys}$ gene clusters in *Drosophila*. J. Mol. Biol., **147:** 475.

Green, C. J., Sohel, I., and Vold, B. S. (1990). The discovery of new intron-containing human tRNA genes using the polymerase chain reaction. J. Biol. Chem., **265:** 12139-12142.

Grunstein, M., and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA, **72:** 3961-3965.

Hall, B. D., Clarkson, S. G., and Tocchini-Valentini, G. (1982). Transcription initiation of eukaryotic transfer RNA genes. Cell, **29:** 3-5.

Hatfield, D., and Rice, M. (1986). Aminoacyl-tRNA (anticodon):codon adaptation in human and rabbit reticulocytes. Biochem. Int., **13:** 835-842.

Hatlen, L., and Attardi, G. (1971). Proportion of the HeLa cell genome complementary to transfer RNA and 5 S RNA. J. Mol. Biol., **56:** 535-553.

Hearst, J., and Stockmayer, W. H. (1962). Sedimentation constants of broken chains and wormlike coils. J. Chem. Phys., **37:** 1425-1433.

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., and Marquisee, M., Merrill, S.H., Penswick, J.R., and Zamir, A. (1965). Structure of a ribonucleic acid. Science, **147:** 1462-1465.

Hu, N.-T., and Messing, J. (1982). The making of strand-specific M13 probes. Gene, **17:** 217-277.

Jacobson, H., and Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. J. Chem. Phys., **18:** 1600-1606.

Jager, D. and Philippsen, P. (1989). Stabilization of dicentric chromosomes in *Saccharomyces cerevisiae* by telomere addition to broken ends or by centromere deletion. EMBO J., **8:** 247-254.

Jeffreys, A. J., Brookfield, J. F. Y., and Semeonov, R. (1985). Positive identification of an immigration test-case using human DNA fingerprints. Nature, **317:** 818-819.

Jelinek, W. R., and Schmid, C. W. (1982). Repetitive sequences in eukaryotic DNA and their expression. Ann. Rev. Biochem., **51:** 813-844.

Johnson, P. F., and Abelson, J. (1983). The yeast $tRNA^{Tyr}$ gene intron is essential for correct modifications of its tRNA product. Nature, **302:** 681-687.

Johnson, G. D., Pirtle, I. L., and Pirtle, R. M. (1985). The nucleotide sequence of tyrosine transfer RNAQ-STAR-PSI-A from bovine liver. Arch. Bioch., **236:** 448-453.

Kassavetis, G. A., Braun, B. R., Nguyen, L. H., and Geiduschek, E. P. (1990). *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. Cell, **60:** 235-245.

Kim, S. H., Sussman, J. L., Suddath, F. L., Quigley, G. J., McPherson, A., Wang A. H. J., Seeman, N. C., and Rich, A. (1974). The general structure of transfer RNA molecules. Proc. Natl. Acad. Sci. USA, **71:** 4970-4974.

Kim, S. H., and Sussman, J. L. (1978). Transfer RNA: structure and function correlation, In Quagliariello, E., Palmieri, F. and Singer, T. P. (Eds). Horizons in Biochemistry and Biophysics. (p 159) New York, Addison-Wesley Publ. Co.

Klekamp, M. S., and Weil, P. A. (1986). Partial purification and characterization of the *Saccharomyces cerevisiae* transcription factor TFIIIB. J. Biol. Chem., **261:** 2819-2827.

Klug, A., Robertus, J. D., Ladner, J. E., Brown, R. S., and Finch, J. T. (1974). Conservation of the molecular structure of yeast phenylalanine transfer RNA in two crystal forms. Proc. Natl. Acad. Sci. USA, **71:** 3711-3715.

Kornberg, R. D. (1974). Chromatin structure - repeating units of histones and DNA. Science, **184:** 868-871.

Krayev, A. S., Markusheva, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Bayev, A. A., and Georgiev, G. P. (1982). Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing. Biochem. Biophys. Res. Commun., **124:** 514-522.

Lassar, A. B., Martin, P. L., and Roeder, R. G. (1983). Transcription of class III genes: formation of preinitiation complexes. Science, **222:** 740-748.

Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.-A., and Böck, A. (1988). Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. Nature, **331:** 723-725.

Lin, V. K. and Agris, P. F. (1980). Alterations in tRNA isoaccepting species during erythroid differentiation of the Friend leukemia cell. Nucl. Acids Res., **8:** 3467-3480.

Lobo, S. M., Ifill, S., and Hernandez, N. (1990). *Cis*-acting elements required for RNA polymerase II and III transcription in the human U2 and U6 snRNA promoters. Nucl. Acids Res., **18:** 2891-2899.

Loh, E. Y., Elliot, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989). Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor d chain. Science, **243:** 217-220.

Ma, D. P., Lund, E., Dahlberg, J. E., and Roe, B. A. (1984). Nucleotide sequences of two regions of the human genome containing $tRNA^{Asn}$ genes. Gene, **28:** 257-262.

MacPherson, J. M. (1988) Isolation and sequence of 6 tyrosine tRNA genes. Ph.D. thesis, University of Alberta.

MacPherson, J. M., and Roy, K. L. (1986). Two human tyrosine tRNA genes contain introns. Gene, **42:** 101-106.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavage. Methods Enzymol., **65:** 499-560.

McAlpine, P. J., Stranc, L. C., Boucheix, C., and Shows, T. B. (1990). The 1990 catalog of mapped genes and report of the nomenclature committee. Cytogenet. Cell. Genet., **55:** 5-76.

McBride, O. W., Melini, R., and Hatfield, D. (1987). Opal suppressor phosphoserine tRNA gene and pseudogene are located on human chromosome 19 and 22, respectively. J. Biol. Chem., **262:** 11163-11166.

McBride, O. W., Pirtle, I. L., and Pirtle, R. M. (1989). Localization of three DNA segments encompassing tRNA genes to human chromosomes 1, 5, and 16: Proposed mechanism and significance of tRNA gene dispersion. Genomics, **5:** 561-573.

Morrison, N., Goddard, J. P., Ledbetter, D. H., Boyd, E., Bourn, D., and Connor, J. M. (1991). Chromosomal assignment of a large tRNA gene cluster ($tRNA^{Leu}$, $tRNA^{Gln}$, $tRNA^{Lys}$, $tRNA^{Arg}$, $tRNA^{Gly}$) to 17p13.1. Hum. Genet., **87:** 226-230.

Morse, R. H. (1989). Nucleosomes inhibit both transcriptional initiation and elongation by RNA polymerase III *in vitro*. EMBO J., **8:** 2343-2351.

Mueller, P. R., and Wold, B. (1989). *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. Science, **246:** 780-786.

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods Enzymol., **155:** 335-350.

Müller, F. and Clarkson, S. G. (1980) Nucleotide sequence of genes coding for $tRNA^{Phe}$ and $tRNA^{Tyr}$ from a repeating unit of *X. laevis* DNA. Cell., **19:** 345-353.

Naylor, S. L., Sakaguchi, A. Y., Shows, T. B., Grzeschik, K.-H., Holmes, M., and Zasloff, M. (1983). Two nonallelic $tRNA_i^{Met}$ genes are located in the p23-q12 region of human chromosome 6. Proc. Natl. Acad. Sci. USA, **80:** 5027-5031.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Rairez-Solis, R., Webster, T. D., Ledbetter, D. H., and Caskey, C. T. (1989). *Alu* polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proc. Natl. Acad. Sci. USA, **86:** 6686-6690.

Normanly, J., and Abelson, J. (1989). tRNA identity. Ann. Rev. Biochem., **58:** 1029-1049.

O'Neill, V. A., Eden, F. C., Pratt, K., and Hatfield, D. L. (1985). A human opal suppressor transfer RNA gene and pseudogene. J. Biol. Chem., **260:** 2501-2508.

Ochman, H., Ajioka, J. W., Garza, D., and Hartl, D. L. (1988). Inverse polymerase chain reaction. In Erlick, H. A. (Ed), PCR Technology: Principles and Applications of DNA Amplification. (p 105). New York, Stockton Press.

Olsen, S., Pirtle, I. L., Pirtle, R. M., and McBride, O. W. (1987). Chromosomal mapping of a cluster of transfer RNA genes to human chromosome 14. Am. J. Hum. Genet., **41:** A 180.

Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B., and Riggs, A. D. (1989). Genomic sequencing and methylation analysis by ligation mediated PCR. Science, **246:** 810-813.

Pieler, T., Appel, B., Oei, S. C., Mentzel, H., and Erdmann, V. A. (1985). Point mutation analysis of the *Xenopus laevis* 5S rRNA gene promoter. EMBO J., **4:** 1847-1853.

Pieler, T., Hamm, J., and Roeder, R. G. (1987). The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. Cell, **48:** 91-100.

Pirtle, I. L., Shortridge, R. D., and Pirtle, R. M. (1986). Nucleotide sequence and transcription of a human glycine $tRNA_{GCC}$ gene and nearby pseudogene. Gene, **43:** 155-167.

Pontius, B. W., and Berg, P. (1991). Rapid renaturation of complementary DNA strands mediated by cationic detergents: A role for high-probability binding domains in enhancing the kinetics of molecular assembly processes. Proc. Natl. Acad. Sci. USA, **88:** 8237-8241.

Pratt, K., Eden, F., You, K. H., O'Neill, V. A., and Hatfield, D. (1985). Conserved sequences in both coding and 5' flanking regions of mammalian opal suppressor tRNA genes. Nucl. Acids Res., **13:** 4765-4775.

Reilly, J. G., Ogden, R., and Rossi, J. J. (1982). Isolation of a mouse pseudo tRNA gene encoding CCA - a possible example of reverse flow of genetic information. Nature, **300:** 287.

Reinberg, D., Horikoshi, M., and Roeder, R. G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II. Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. J. Biol. Chem., **262:** 3322-3330.

Rigaud, G., Grange, T., and Pictet, R. (1987). The use of NaOH as transfer solution of DNA onto nylon membranes decreases the hybridization efficiency. Nucl. Acids Res., **15:** 857.

Rinehart, F. P., Ritch, T. G., Deininger, P. L., and Schmid, C. W. (1981). Renaturation rate studies of a single family of interspersed repeated sequences in human deoxyribonucleic acid. Biochemistry, **20:** 3003-3010.

Rogers, J. H. (1985). The origin and evolution of retroposons. Int. Rev. Cytol., **93:** 187-279.

Rosenthal, A., and Jones, D. S. C. (1990). Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. Nucl. Acids Res., **18:** 3095.

Rould, M. A., Perona, J. J., Söll, D., and Steitz, T. A. (1989). Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8Å resolution. Science, **246:** 1135-1142.

Rould, M. A., Perona, J. J., and Steitz, T. A. (1991). Structural bases of anticodon loop recognition by glutaminyl-tRNA synthetase. Nature, **352:** 213-218.

Roy, K. L., Cooke, H., and Buckland, R. (1982). Nucleotide sequence of a segment of human DNA containing three tRNA genes. Nucl. Acids. Res., **10:** 7313-7322.

Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T., and Schmid, C. W. (1980). Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature, **284:** 372-374.

Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991). Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. Science, **252:** 1682-1689.

Saiki, R. K., Gelfand, S., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, **239:** 487-494.

Sakamoto, K., and Okada, N. (1985). Rodent type 2 Alu family, rat identifier sequence, rabbit C family, and bovine or goat 73-bp repeat may have evolved from tRNA genes. J. Mol. Evol., **22:** 134-140.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, **74:** 5463-5467.

Santos, T., and Zasloff, M. (1981). Comparative analysis of human chromosomal segments bearing nonallelic dispersed $tRNA_i^{Met}$ genes. Cell, **23:** 699-709.

Scheller, R. H., Anderson, D. M., Posakony, J. W., McAllister, L. B., Britten, R. J., and Davidson, E. H. (1981). Repetitive sequences of the sea urchin genome. J. Mol. Biol., **149:** 15-39.

Schimmel, P. R. (1987). Aminoacyl tRNA synthetases: General scheme of structure-function relationships in the polypeptides and recognition of transfer RNAs. Ann. Rev. Biochem., **56:** 125-158.

Schimmel, P. R., and Söll, D. (1979). Aminoacyl-tRNA synthetases: General features and recognition of transfer RNAs. Ann. Rev. Biochem., **48:** 601-648.

Schmid, C. W., and Deininger, P. L. (1975). Sequence organization of the human genome. Cell, **6:** 345-358.

Schulman, L. H., and Abelson, J. (1988). Recent excitement in understanding transfer RNA identity. Science, **240:** 1591-1592.

Segall, J. (1986). Assembly of a yeast 5S RNA gene transcription complex. J. Biol. Chem., **261:** 11578-11584.

Sharp, S., DeFranco, D., Dingermann, T. Farrell, P. and Söll, D. (1981). Internal control regions for transcription of eukaryotic transfer RNA genes. Proc. Natl. Acad. Sci. USA, **78:** 6657-6661.

Sharp, S. J., Schaak, J., Cooley, L., Burke, D. J., and Söll, D. (1985). Structure and transcription of eukaryotic tRNA genes. CRC Crit. Rev. Bioch., **19:** 107-144.

Shibuya, K., Noguchi, S., Nishimura, S., and Sekiya, T. (1982). Characterization of a rat tRNA gene cluster containing the genes for $tRNA^{Asp}$, $tRNA^{Gly}$, and $tRNA^{Glu}$, and pseudogenes. Nucl. Acids Res., **10:** 441.

Shortridge, R. D., Johnson, G. D., Craig, L. C., Pirtle, I. L., and Pirtle, R. M. (1989). A human tRNA gene heterocluster encoding threonine, proline, valine tRNAs. Gene, **79:** 309-324.

Shortridge, R. D., Pirtle, I. L., and Pirtle, R. M. (1985). Nucleotide sequence and transcription of a gene coding human $tRNA^{Gly}_{CCC}$. Gene, **33:** 269-277.

Sollner-Webb, B., and Tower, J. (1986). Transcription of cloned eukaryotic ribosomal tRNA genes. Ann. Rev. Biochem., **55:** 801-830.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., **98:** 503-517.

Sprinzl, M., Hartmann, T., Meissner, F., Moll, J., and Vorderwulbecke, T. (1987). Compilation of tRNA sequences and sequences of tRNA genes. Nucl. Acids Res., **15:** r53-r188.

Swanson, R., Hoben, P., Sumner-Smith, M., Uemura, H., Watson, L., and Söll, D. (1988). Accuracy of *in vivo* aminoacylation requires proper balance of tRNA and aminoacyl tRNA synthetase. Science, **242:** 1548-1551.

Temple, G. F., Dozy, A. M., Roy, K. L., and Kan, Y. W. (1982). Construction of a functional human suppressor tRNA gene: an approach to gene therapy for β-thalassaemia. Nature, **296:** 537-540.

Tindall, K. R., and Kunkel, T. A. (1988). Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. Biochemistry, **27:** 6008-6015.

Triglia, T., Peterson, M. G., and Kemp, D. J. (1988). A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucl. Acids Res., **16:** 8186.

van Tol, H., and Beier, H. (1988). All human $tRNA^{Tyr}$ genes contain introns as a prerequisite for pseudouridine biosynthesis in the anticodon. Nucl. Acids Res., **16:** 1951-1966.

van Tol, H., Stange, N., Gross, H. J., and Beier, H. (1987). Human and a plant intron-containing $tRNA^{Tyr}$ gene are both transcribed in a HeLa cell extract but spliced along different pathways. EMBO, **6:** 35-41.

Vanin, E. F. (1985). Processed pseudogenes: Characteristics and evolution. Ann. Rev. Genet., **19:** 253-272.

Vassart, G., Georges, M., Monsieur, R., Brocas, H., LeQuarre, A.S., and Christopher, D. (1987). A sequence in M13 phage detects hypervariable minisatellites in human and animal DNA. Science, **235**: 683-684.

Vieira, J., and Messing, J. (1987). Production of single-stranded plasmid DNA. Methods Enzymol., **153:** 3-11.

Vosberg, H.-P. (1989). The polymerase chain reaction: an improved method for the analysis of nucleic acids. Hum. Genet., **83:** 1-15.

Waldschmidt, R., Jahn, D., and Seifart, K. H. (1988). Purification of transcription factor IIIB from HeLa cells. J. Biol. Chem., **263:** 13350-13356.

Wang, C. K., and Weil, P. A. (1989). Purification and characterization of *Saccharomyces cerevisiae* transcription factor IIIA. J. Biol. Chem., **264:** 1092-1099.

Wang, J. C., and Davidson, N. (1966). On the probability of ring closure of lambda DNA. J. Mol. Biol., **19:** 469-482.

Weiner, A. M., Deininger, P. L., and Efstratiadis, A. (1986). Nonviral retroposons: Genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Ann. Rev. Biochem., **55:** 631-661.

Wetmur, J. G., and Davidson, N. (1968). Kinetics of renaturation of DNA. J. Mol. Biol., **31:** 349-370.

Wolfe, A. P., and Brown, D. D. (1988). Developmental regulation of two 5S ribosomal RNA genes. Science, **241:** 1626-1632.